Development of a case database including atypical cases of rare diseases with emphasis on symptoms and findings using generative AI

Eisuke Dohi¹), Itaru Hayakawa²), Tomoyasu Matsubara³), Toyofumi Fujiwara⁴), Tomoyasu Yamamoto⁴)

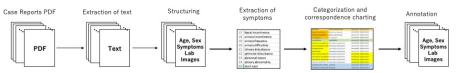
National Center of Neurology and Psychiatry, National Institute of Neuroscience, 2) National Center for Child Health and Development, Department of 1) Neurology 3) Tokushima University Department of Neurology, 4) Database Center for Life Science

Background

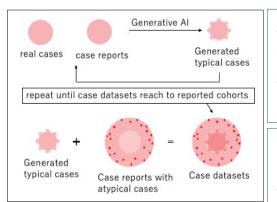
In the case of intractable and rare diseases, the sheer number of diseases and the small number of individual cases prevents medical practitioners from gaining sufficient case experience. Educational resources are also limited to textbooks with typical cases and case reports biased toward atypical cases, and there is a lack of educational resources to grasp the overall picture of the disease. Even if all of the published learning resources were collected, case reports are biased toward atypical cases, making it difficult to create a patient distribution that reflects the image of the disease seen in actual clinical practice.

Therefore, this study aims to create data resources for learning and diagnostic support by constructing a case database that is similar to real clinical cases using generative AI.

Method



1: Case data were extracted from case reports PDF and symptoms and findings were categorized. Case data were annotated with corresponding charts. All the procedure were conducted with generative AI (LLM such as ChatGPT) except making the corresponding charts.



2: To compensate typical cases, typical cases will be generated with generative AI with typical cases in real cases and case reports. The additional generated typical cases will be added to case reports datasets until basic statics reach previously reported cohort studies.

3: Case dataset will be visualized based on their symptoms and findings in individual cases to ease to understand the atypical case among the case datasets.

Results

(1) PDF \rightarrow text extraction : Ask your PDF pro works well! 2 The extraction of symptoms : LLM works well (+/-, time course)

You 下記の文章から症状を抜き出してください 現病歴1歳7カ 月頃までは順調に発育していたが,1歳9カ月頃より特異な手の動きが出 てきた.はじめは手を開いて胸の前でうちあわせる動きであったが次第に手指を揃 え,軽く第2指間関節で曲げた形で右手掌に左手をのせる様にたたく。そして時々左 指を口へ持ってゆくという動きが出てきた.この頃より今までできていた,ものをつ かむ つまむ 指差し等しなくなり いたずらも減り 言葉も出なくなり 音志表示が不能 瞭になってきた.母親との視線も合わなくなり,周囲への関心も減り,表情が乏しくな り,笑わなくなってきた.種々病院を受診し相談するも,脳波に異常所見(左優位に棘波 頻発)が認められた以外、血液・尿初見は全て正常範囲であった。3歳4ヶ月紹介さ れ当クリニックを受診する。



症状•所見	有無	時点		
発達遅延と進行性の感怒拡大("probable hydrocephalus")	3.19	4ヶ月間		
神経学的疾患の家族歴の欠如	なし	8020-09		
現图の増加(25hから80ibパーセンタイル)	3,17	80.021-019		
身長と体重がJardパーセンタイル以下	2,17	(1) (1) (1) (1) (1) (1) (1) (1) (1) (1)		
CTとMRIでの前頭集白質の減少	3,19	802040		
CTとMRIでの水道症や腫瘍の否定	3,17	初始時		
腰椎穿刺の開放圧(304mm H2O)	3,17	初診時の腰椎穿刺		
腰椎穿刺の閉鎖圧(154mm H2O)	2,17	初辺時の便検定利		
腰椎穿刺の結果(蛋白質240 mg/L, グルコース3.3 mmel/L)	3,19	初設時の優権穿刺		
デジタル脳血管撮影での静脈洞閉塞の証拠なし	3,17	初診後の評価時		
甲状腺機器、血中能レベル、酵素アッセイでの正常所見	3,17	初診後の評価時		
脳波での後期部中心の軽度の一般化遅延	2,17	初節後の評価時		
腰椎穿利の繰り返し(開放)E110mm H2O, 閉鎖)E80 mm H2O)	2,19	最初の腰椎穿刺から3日後		
腰椎穿刺の袖9窓し(開放圧200mm H2O, 閉鎖圧160 mm H2O)	3.17	最初の課程室利から1週間目		

3 Annotation with LLM did not work with HPO....

Annotation with corresponding chart works! <u>visualization pf typical</u>								
<u>corre</u>	spondin	ig chart		Annotated clinical text	and atypical cases			
enurests	secal incarditence	autonomic dystunction		and deteriorating cognition <u><"demetia"></u> , He was born normally and achieved	-8-			
	faecal incontinence	autonomic dysfunction		developmental milestones without delay. The				
	constipation	autonomic dysfunction	LLM		-26 - 5			
constipation	constipation	autonomic dysfunction		family history was noncontributory. At the age				
episodic vasovagal sync	orthostatic hypotension	autonomic dysfunction		of 2, he had 2 simple febrile seizure<"seizure">s. At the age of 8, he had a	-m- 5 . get			
	orthostatic hypotension	autonomic dysfunction			2 1 1 S 1			
	arthostatic hypotension	autonomic dysfunction		generalized febrile convulsion that lasted				
	orthostatic hypotension	autonomic dysfunction		about 10 min. Subsequent				
	arthostatic hypotension	autonomic dysfunction	-	electroencephalography (EEG) revealed	10 10 10 10 10 10 10 10 10 10 10 10 10 1			
orthostatic hypotentic	orthostatic hypotension orthostatic hypotension	autonomic dysfunction autonomic dysfunction		bilateral frontal spikes during sleep.				

Current issues and solusions

(1) It is unclear if the HPO does not have term or if the annotation function is inadequate.

epilepsy. Since then, his school performance slowly began to deteriorate

(2) Streamlining the creation of correspondence tables.

(3) Stability o results when using LLM

Organizing the HPO to make it easier for clinicians to understand. Automate each procedure and clarify where expert help is needed.

Acknowledgments and Fundings

Special thanks to Yuka Tateishi, Terue Takatsuki and our mentors, advisors, and colleagues for their guidance and support.

Kakenhi 22K12253, 23K11886, Rois-DS Joint 2023