

「総括」

1. 登録された症例数は、平成9年度と比較すると飛躍的に増加したが、従来厚生省や母子愛育会が報告してきた症例数よりは少ない数であり、十分に網羅されていないと考えられた。
2. 全体を通じて医療意見書への記入については不十分であり、無記入欄が非常に多かったため、これを改める方法を検討するべきと思われた。費用－便益の立場からも、それによってスクリーニングの効果が判定できるようにするべきであると考えられた。
3. 記入項目については有無のみであり、その内容についての記載が必要と考えられることも多かった。例えば先天性代謝異常症の意見書では、合併症は有無のみの記入であり、どのような内容の合併症であるか判断が不能であった。
4. クレチン症については ICD の変更を求める意見があり、若干の変更が必要と考えられた。
5. 新生児マス・スクリーニングで発見されてくる疾患は症状発現がないように十分な治療が求められている。治療の内容、特に食事療法についての記載が必要であり、治療成績はこれから発見されてくる秋冷の治療を一層向上することに役立つと考えられる。医療意見書を手がかりにして、症例を治療している医師に対して治療経過を求めることが必要になる場合もある。国の事業として実施されているスクリーニングは、それにより利益を最大限にすることが求められており、同時にスクリーニングにより得られる利益は個人に還元されるべきであり、そのため医療意見書の研究利用が求められる。

表1 PKUの年齢別、TYPE別、発見別分類

年齢	症例数	新規	転入	継続	無記入	スクリーニング	無記入	空欄
0歳	18	8	0	7	3	13	5	0
1歳	5	0	0	5	0	5	0	0
2歳	11	1	0	9	1	10	1	0
3歳	5	1	0	3	1	5	0	0
4歳	5	0	0	5	0	4	1	0
5歳	5	0	0	4	1	3	2	0
6歳	6	1	0	5	0	4	2	0
7歳	6	0	0	3	3	4	2	0
8歳	4	0	0	4	0	2	2	0
9歳	6	0	0	5	1	5	1	0
10歳	7	0	0	5	2	5	2	0
11歳	4	0	0	2	2	3	1	0
12歳	4	0	0	4	0	1	3	0
13歳	5	1	0	4	0	4	1	0
14歳	8	1	0	7	0	6	2	0
15歳	4	0	0	4	0	4	0	0
16歳	4	0	0	4	0	2	2	0
17歳	6	0	0	5	1	6	0	0
18歳	0	0	0	0	0	0	0	0
19歳	1	0	0	1	0	1	0	0
無記入	12	1	0	11	0	4	8	0
合計	126	14	0	97	15	91	35	0

表2 MSUDの年齢別、TYPE別、発見別分類

年齢	症例数	新規	転入	継続	無記入	スクリーニング	無記入	空欄
0歳	3	3	0	0	0	2	1	0
1歳	0	0	0	0	0	0	0	0
2歳	1	0	0	1	0	0	1	0
3歳	0	0	0	0	0	0	0	0
4歳	2	0	0	2	0	2	1	0
5歳	1	0	0	1	0	0	0	0
6歳	0	0	0	0	0	0	0	0
7歳	1	0	0	1	0	1	0	0
8歳	1	0	0	1	0	1	0	0
9歳	3	0	0	2	1	3	0	0
10歳	1	1	0	0	0	1	0	0
11歳	1	0	0	1	0	0	0	0
12歳	2	0	0	2	0	2	0	0
13歳	0	0	0	0	0	0	0	0
14歳	2	0	0	2	0	1	1	0
15歳	3	0	0	3	0	1	2	0
16歳	1	0	0	1	0	0	2	0
17歳	0	0	0	0	0	0	0	0
18歳	0	0	0	0	0	0	0	0
19歳	0	0	0	0	0	0	0	0
無記入	1	0	0	1	0	0	1	0
合計	23	4	0	18	1	14	9	0

表3 ホモシステン尿症の年齢別、TYPE別、発見別分類

年齢	症例数	新規	転入	継続	無記入	スクリーニング	無記入	空欄
0歳	1	1	0	0	0	1	0	0
1歳	1	0	0	1	0	0	1	0
2歳	2	0	0	1	1	2	0	0
3歳	0	0	0	0	0	0	0	0
4歳	1	0	0	1	0	1	0	0
5歳	2	0	0	1	1	2	0	0
6歳	3	0	0	2	1	3	0	0
7歳	0	0	0	0	0	0	0	0
8歳	2	0	0	1	1	1	1	0
9歳	0	0	0	0	0	0	0	0
10歳	1	0	0	1	0	1	0	0
11歳	2	0	0	2	0	1	1	0
12歳	2	0	0	2	0	0	2	0
13歳	0	0	0	0	0	0	0	0
14歳	3	0	0	2	1	0	3	0
15歳	1	0	0	1	0	0	1	0
16歳	0	0	0	0	0	0	0	0
17歳	2	0	0	1	1	2	0	0
18歳	0	0	0	0	0	0	0	0
19歳	0	0	0	0	0	0	0	0
無記入	2	1	0	1	0	2	1	0
合計	25	2	0	17	6	16	9	0

表4 ガラクトース血症年齢別、TYPE別、発見別分類

年齢	症例数	新規	転入	継続	無記入	スクリーニング	無記入	空欄
0歳	10	8	0	1	1	7	3	0
1歳	12	1	0	8	3	11	1	0
2歳	8	2	0	6	0	4	4	0
3歳	6	2	0	4	0	5	1	0
4歳	9	0	0	8	1	3	5	1
5歳	2	0	0	2	0	2	0	0
6歳	4	0	0	4	0	2	1	1
7歳	3	0	0	2	1	2	1	0
8歳	2	1	0	1	0	2	0	0
9歳	2	0	0	2	0	1	1	0
10歳	2	0	0	2	0	2	0	0
11歳	0	0	0	0	0	0	0	0
12歳	3	0	0	3	0	0	3	0
13歳	2	0	0	2	0	2	0	0
14歳	0	0	0	0	0	0	0	0
15歳	0	0	0	0	0	0	0	0
16歳	1	0	0	1	0	0	1	0
17歳	3	0	0	3	0	2	1	0
18歳	0	0	0	0	0	0	0	0
19歳	0	0	0	0	0	0	0	0
無記入	3	0	0	3	0	2	1	0
合計	72	14	0	52	6	47	23	2

表5 調査別症例数の比較

	小樽	厚生省	愛育会
フェニルケトン尿症	126	371	242
メープルシロップ尿症	23	59	42
ホモシスチン尿症	25	158	27
ガラクトース血症	72	802	265

表6 先天性代謝異常症4疾患の情況

	フェニルケトン尿症	メープルシロップ尿症	ホモシスチン尿症	ガラクトース血症
スクリーニング	90	14	16	47
無記入	32	8	9	21
空欄	4	1	0	4
知能障害あり	3(2.40%)	4(17.40%)	5(20.00%)	1(1.40%)
無記入	6	19	20	67
空欄	217	0	0	4
痙攣あり	1(0.79%)	3(13.00%)	5(20.00%)	1(1.40%)
無記入	219	20	20	67
空欄	6	0	0	4
就学情況				
通常就学	55	8	11	18
障害児学級	1(0.78%)	0	1(4.00%)	1
養護学級	0	4(17.39%)	3(12.00%)	1
訪問教育	0	0	0	0
その他	3	2	0	2
無記入	61	9	10	47
空欄	6	0	0	3
経過				
治癒	0	0	0	1
寛解	13	5	2	18
改善	58	4	7	21
不変	29	9	11	14
再燃	0	0	0	0
悪化	0	0	0	0
死亡	0	0	0	0
判定不能	0	0	0	1
その他	22	5	5	16
空欄	4	0	0	1
症例数	126	23	25	72

表7 フェニルケトン尿症地域別報告数

番号	地域	報告数	番号	地域	報告数	番号	地域	報告数	番号	地域	報告数
1	北海道		21	岐阜	1	41	佐賀		61	郡山市	
2	青森		22	静岡	4	42	長崎		62	宇都宮市	
3	岩手	3	23	愛知	2	43	熊本	3	63	新潟市	1
4	宮城	5	24	三重	3	44	大分		64	富山市	2
5	秋田	3	25	滋賀		45	宮崎	4	65	金沢市	1
6	山形		26	京都	4	46	鹿児島		66	岐阜市	1
7	福島		27	大阪	8	47	沖縄	1	67	静岡市	
8	茨城	2	28	神戸		48	札幌市	5	68	浜松市	1
9	栃木		29	奈良	2	49	仙台市		69	豊田市	
10	群馬	1	30	和歌山	5	50	千葉市		70	堺市	1
11	埼玉		31	鳥取		51	横浜市		71	姫路市	
12	千葉	2	32	島根		52	川崎市		72	和歌山市	3
13	東京	21	33	岡山	1	53	名古屋市	6	73	岡山市	3
14	神奈川	3	34	広島	2	54	京都市		74	福山市	3
15	新潟	1	35	山口		55	大阪市		75	高知市	1
16	富山	2	36	徳島	3	56	神戸市		76	長崎市	
17	石川		37	香川	2	57	広島市		77	熊本市	2
18	福井	1	38	愛媛		58	北九州市	6	78	大分市	
19	山梨		39	高知		59	福岡市		79	宮崎市	
20	長野		40	福岡		60	秋田市		80	鹿児島市	2

表8-1 クレチン症の地域別報告数

	都道府県市	小児慢性H10年度				愛育会H9年度新規				厚生省 H10年度
		総数	総0歳	新規総数	新規0歳	総数	クレチン	TH	THTGH	
1	北海道									12
2	札幌市	48	1	2	1					6
3	青森県					9	5	0	4	4
4	岩手県	55	5	12	4	11	10	0	1	2
5	宮城県	51	3	4	1	2	2	0	0	1
6	仙台市					3	0	2	1	6
7	秋田県	7	0	0	0	5	2	0	3	7
8	秋田市									
9	山形県					8	4	2	2	9
10	福島県					7	2	2	3	5
11	郡山市									7
12	茨城県	44	6	9	6	3	2	0	1	
13	栃木県					5	5	0	0	4
14	宇都宮市	8	0	0	0					
15	群馬県	3	3	1	1	16	5	3	8	8
16	埼玉県					7	6	1	0	9
17	千葉県	59	8	10	6	33	22	4	7	5
18	千葉市	18	3	4	2					2
19	東京都	214	24	19	10	8	3	2	3	45
20	神奈川県	34	6	10	4					19
21	横浜市									9
22	川崎市									0
23	新潟県	56	12	11	11					12
24	新潟市	5	5	5	5					
25	富山県	36	5	6	3	10	1	2	7	6
26	富山市	14	1	2	1					
27	石川県	2	0	0	0	5	5	0	0	3
28	金沢市	9	1	0	0					
29	福井県	17	1	1	0	3	3	0	0	4
30	山梨県	14	2	3	2	2	0	0	2	4
31	長野県					10	6	1	3	5
32	岐阜県	29	1	1	0					11
33	岐阜市	0	0	0	0					
34	静岡県	113	14	10	6	10	10	0	0	18
35	静岡市									
36	浜松市	16	2	3	2					
37	愛知県	4	0	0	0					20
38	豊田市									
39	名古屋市	46	2	9	1	19	11	4	4	12
40	三重県	37	1	3	1					0
41	滋賀県					15	14	0	1	10
42	京都府	63	8	15	7	10	5	0	5	9
43	京都市									52

表8-2 クレチン症の地域別報告数

44	大阪府	186	21	29	19					23
45	堺市	27	3	8	2					
46	大阪市					12	7	1	4	5
47	兵庫県					44	10	12	22	13
48	姫路市									
49	神戸市	5	0	2	0	20	6	6	8	8
50	奈良県	18	0	2	0	19	10	1	8	6
51	和歌山県	16	2	1	1	7	3	0	4	7
52	和歌山市	6	1	0	0					
53	鳥取県									0
54	鳥根県					2	0	0	2	6
55	岡山県	33	5	11	5	3	3	0	0	10
56	岡山市	22	2	1	1					
57	広島県	23	5	7	4					8
58	福山市	2	1	2	1					
59	広島市	16	6	7	5					15
60	山口県	11	0	1	0	5	2	1	2	4
61	徳島県	18	0	4	0	6	4	2	0	2
62	香川県	9	1	1	1	7	2	2	3	0
63	愛媛県	27	4	4	3	11	6	0	5	5
64	高知県	13	2	3	2	20	3	6	11	3
65	高知市	5	0	0	0					
66	福岡県					46	19	10	17	6
67	北九州市	15	2	6	2					6
68	福岡市									14
69	佐賀県	2	0	1	0	1	0	0	1	1
70	長崎県	0	0	0	0	6	0	3	3	8
71	長崎市	2	1	0	0					
72	熊本県	55	9	2	2	2	1	0	1	10
73	熊本市	2	2	2	2					
74	大分県	30	4	9	2	9	5	0	4	4
75	大分市	0	0	0	0					
76	宮崎県	5	3	3	3	8	2	1	5	9
77	宮崎市									
78	鹿児島県	13	10	9	8	20	8	0	12	14
79	鹿児島市	1	0	0	0					
80	沖縄県	63	11	16	8	14	14	0	0	14

表9 集計別クレチン症発生頻度(何人に1人)

都道府県市	検査総数	小慢	愛育会	厚生省	都道府県	検査総数	小慢	愛育会	厚生省
北海道	32054			2671	京都府	9558	1195	1912	1062
札幌市	17608	17608		2935	京都市	15203			292
青森県	15106		3021	3777	大阪府	66724	2780		2901
岩手県	13807	2761	1381	6904	堺市				
宮城県	11343	3781	5672	11343	大阪市	28003		4000	5601
仙台市	10875			1813	兵庫県	40456		4046	3112
秋田県	10157		5097	1451	姫路市				
秋田市					神戸市	13908		2318	1739
山形県	11993		2998	1333	奈良県	13455		1346	2243
福島県	22907		11454	4581	和歌山県	10792	3597	3597	1542
郡山市					和歌山市				
茨城県	28798		14339	4114	鳥取県	6251			
栃木県	19407		3881	4852	島根県	7055			1176
宇都宮市					岡山県	19406	2772	6469	1941
群馬県	20855	6952	4171	2607	岡山市				
埼玉県	61313		10219	6813	広島県	10695	1738		713
千葉県	44877	5610	2040	8975	福山市				
千葉市	7392	2464		3696	広島市	14635		7318	3659
東京都	99411	4142	33137	2209	山口県	14635		7318	3659
神奈川県	35054	5842		1845	徳島県	7786		1947	3893
横浜市	29823			3314	香川県	9858	9858	4929	
川崎市	11850				愛媛県	14766	3692	2461	2953
新潟県	24585	1446		2049	高知県	6818	3409	2273	2273
新潟市					高知市				
富山県	10917	1820	10917	1820	福岡県	25310		1332	4218
富山市					北九州市	11411	5706		1902
石川県	12182	12182	2436	4061	福岡市	12765			912
金沢市					佐賀県	9076			9076
福井県	8859	8895	2953	2215	長崎県	15986	15986		1998
山梨県	8612	4306		2153	長崎市				
長野県	22675		3779	4535	熊本県	19022	1729	19022	1902
岐阜県	22128	22128		2012	熊本市				
岐阜市					大分県	11994	2999	2399	2999
静岡県	38202		3820	2122	大分市				
静岡市					宮崎県	13054		6527	1450
浜松市					宮崎市				
愛知県	53515			2676	鹿児島県	17272	1727	2159	1234
豊田市					鹿児島市				
名古屋市	22654	11327	2059	1888	沖縄県	17683	1608	1263	1263
三重県	18766	18766			母数		698065	884828	1E+06
滋賀県	13651		975	1365	患者数		183	228	527
					頻度(1/人数)		3815	3881	2331

表10 CAHの年齢別、TYPE別、発見別分類

	TYPE別				発見別		
	症例数	新規	継続	無記入	スクリーニング	その他	無記入
0歳	48	31	15	2	31	9	8
1歳	31	3	22	6	18	3	10
2歳	28	6	17	2	16	5	7
3歳	27	3	21	3	16	5	6
4歳	31	2	27	2	20	3	8
5歳	34	3	1	27	15	6	13
6歳	30	0	29	1	19	7	4
7歳	35	0	6	28	18	10	7
8歳	36	3	31	2	14	8	14
9歳	22	0	21	1	9	8	5
10歳	33	2	30	1	4	11	17
11歳	26	2	23	1	2	10	14
12歳	26	0	23	2	4	13	9
13歳	34	0	32	2	3	17	14
14歳	32	0	29	1	4	19	9
15歳	35	0	34	1	4	23	8
16歳	32	3	26	3	3	18	11
17歳	20	2	18	0	2	12	6
18歳	5	0	5	0	0	1	4
19歳	1	0	1	0	0	1	0
23歳	1	0	1	0	0	0	1
空欄	41	6	34	1	2	7	31
合計	608	66	446	86	204	196	206

表11 CAH合併症

合併症なし	410
合併症あり	84
無記入	94
空欄	20

表12 CAHの経過別分類

治癒	0
寛解	87
改善	319
不変	80
再燃	0
悪化	1
死亡	0
判定不能	4
その他	95
空欄	22

表13
母子愛育会でのCAH症例数

年度	症例数	死亡数
1991	43	2
1992	27	
1993	47	1
1994	37	
1995	45	1
1996	26	
1997	40	
合計	266	4

表14 先天性副腎皮質過形成症ICD別集計

年齢	E25.0	E25.0A	E25.9	合計
0	45	2	1	48
1	24	3	1	28
2	21	1	5	27
3	25	0	1	26
4	26	4	1	31
5	26	2	5	33
6	28	1	1	30
7	31	1	3	35
8	28	3	4	35
9	19	1	3	23
10	26	1	6	33
11	19	2	5	26
12	18	1	7	26
13	29	1	5	35
14	24	1	6	31
15	26	3	4	33
16	24	0	8	32
17	13	0	5	18
18	2	0	3	5
19	1	0	0	1
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	1	0	0	1
無記入	28	0	11	39
合計	484	27	85	596

厚生科学研究費補助金（子ども家庭総合研究事業）

分担研究報告書

母子保健情報の登録・評価に関する研究

「心身障害研究・子ども家庭総合研究報告書」のデータベース化に関する研究

分担研究者：中村 敬（日本子ども家庭総合研究所 研究企画情報部）

研究協力者：斉藤 進（日本子ども家庭総合研究所 母子保健研究部）

庄司順一（日本子ども家庭総合研究所 研究企画情報部）

恒次欽也（愛知教育大学教育学部）

横田千秋（三井化学株式会社）

モニター：作田亮一（獨協医科大学越谷病院小児科）

瀧川逸朗（東京都立大塚病院小児科）

長坂典子（母子愛育会総合母子保健センター研修部）

上石晶子（東京都母子保健サービスセンター）

松浦賢長（京都教育大学）

【研究要旨】

心身障害研究および厚生科学研究子ども家庭総合研究はわが国の母子医療・保健・福祉に関するトップレベルの研究として、その成果は施策や研究・教育に活用されている。本研究ではこの研究事業の成果を記した研究報告書を永久保存することと、これらの成果を広く普及することを目的として、これらの報告書の電子化およびデータベース化の方法を検討した。研究方法は、すでに印刷された報告書が対象になる過去分の「心身障害研究」報告書と、各研究者から報告書を電子データとして収集可能な平成10年度以降の「子ども家庭総合研究」報告書の2分野に分けて検討し、平成10年度にCD化の方法論を確立した。

今年度は、平成10年度の研究の成果を踏まえ、昭和50年～平成9年（23年分）の過去分「心身障害研究」報告書を解体し1ページごとに画像データとして、イメージスキャナで取り込み電子データ化を図った。また、新規の平成10年度の子どもの家庭総合研究報告書については、研究者から文字情報を生かせる形式で、直接電子データで収集し、データベース化を図った。

【見出語】

厚生省心身障害研究 厚生科学研究 子ども家庭総合研究 報告書 データベース
CDROM 画像データ acrobat

A. 研究目的

過去27年に及ぶ「心身障害研究」および平成10年度から実施された厚生科学研究「子ども家庭総合研究」の成果が広く有効に活用されるためには、より多くの関係者に報告書

が配布され、研究内容が手軽に入
手できる体制を整える必要がある。
このためには、報告書を電子化し、
検索機能を備えたデータベースと
して構築することが最も有効な方
法と考えられる。電子化された報
告書は、従来の印刷物に比べ、コ
ンパクトな媒体に大量のデータを
収納することが可能であり、また、
作成方法を工夫すれば、強力な検
索機能を備えることができる。

本研究班に課せられた目的は、
過去の心身障害研究報告書をデー
タベースとして、実用化するとと
もに、これからの「子ども家庭総
合研究」報告書をデータベース化
するための技術的方法論を確立し、
さらに、その方法に基づいて、実
際に報告書のCD化を図る事であ
る。

B. 研究方法

研究の対象は、1) 過去の心身
障害研究報告書（平成9年度まで）
と、2) 新規の子ども家庭総合研
究報告書（平成10年度以降）の2
つであり、前者はすでに製本され
ている報告書であり、後者は各研
究者から電子データとして収集可
能な報告書である。

研究の方法は、1) に関しては、
平成7～8年度心身障害研究「保
健・医療・福祉にかかわる情報と
社会資源の有効活用に関する研
究」、および平成10年度本研究に
おいて、電子化に関する方法論を

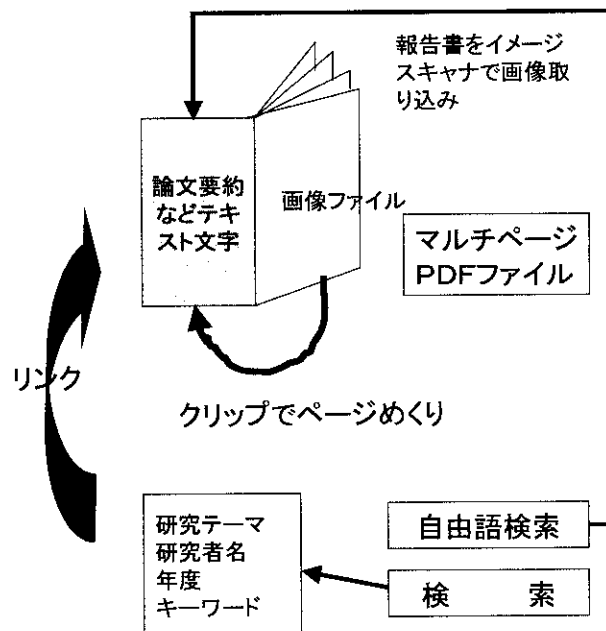
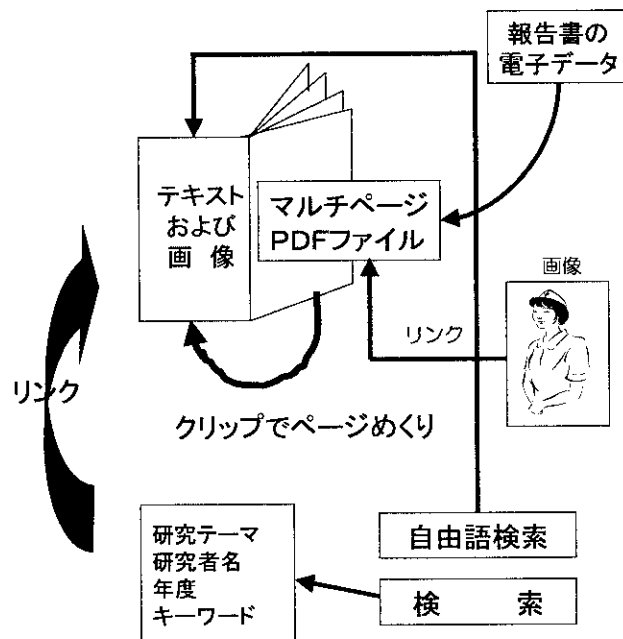


図1：過去分報告書データベース構造



新規分報告書のデータベース構造の基本形
これに、一部画像ファイルで収録したデー
タも混在する形式。

図2：新規分報告書データベース構造

確立した。2) に関しては、平成 10 年度本研究において、現在普及している電子化方法と検索ソフトについて検討し、報告書の電子データの収集方法とデータベース化の方法論を確立した。

本年度は、平成 10 年度で確立した方法論（図 1、図 2）に基づき、

1) 過去の心身障害研究報告書のうち、イメージスキャナで画像として取り込むことが可能な報告書について、1 ページごとにスキャナで取り込み、マルチページ化して電子データ化を図った。昭和 46～49 年の報告書は、製本化し保存してあり、長い年月を経て、紙質が劣化し、変色したり、中には手書きの原稿もあり、画像データとして取り込んでも不鮮明で判読不能な個所が多くなると考えられ、電子化の対象から外した。したがって、本年度の作業は、昭和 50 年から平成 9 年までの 23 年分の報告書について、画像ファイルとして電子化を図った。

2) 新規分報告書（平成 10 年度子ども家庭総合研究報告書）については、報告書と一緒に収集した電子データをもとにデータベースを作成した。

C. 研究結果

本年度作成したデータベースについて、過去の心身障害研究報告書と、新規分子ども家庭総合研究報告書の各データベースに分けて紹介する。

1) データベース検索ソフトのインストールについて（図 3）、

データベース検索ソフトは Alchemy release 6 を用いている。対応機種は IBM/PC 互換機、PC-9800 メモリー16MB 以上（64MB 以上推奨）であり、OS は Windows95、98、NT 4.0 のいずれか、マルチセッション対応の CD-ROM ドライブ、ハードディスク空き容量は 25MB 以上を必要とする。また、アクロバット・リーダーをプラグインしておく必要がある。現時点では Macintosh OS には対応していない。

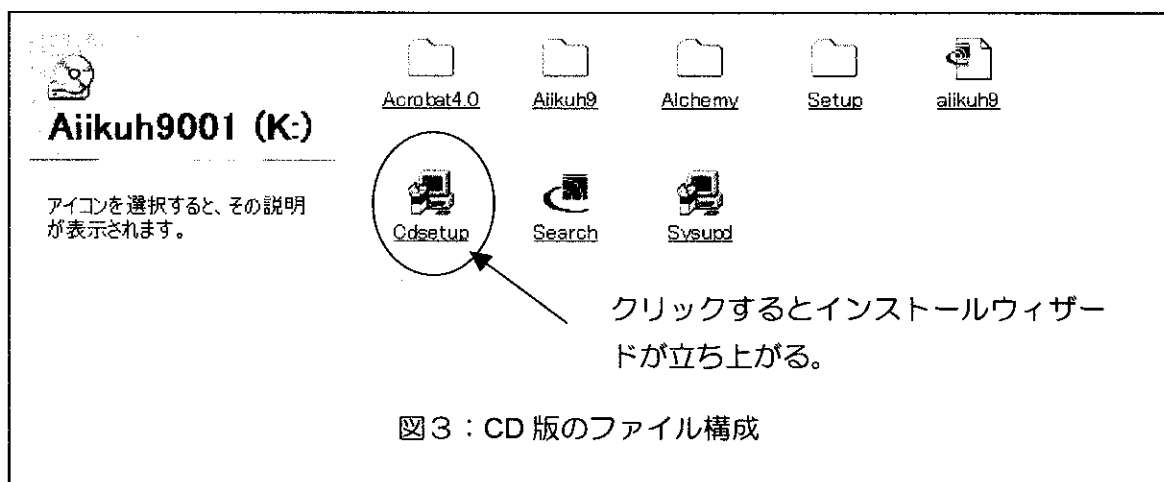
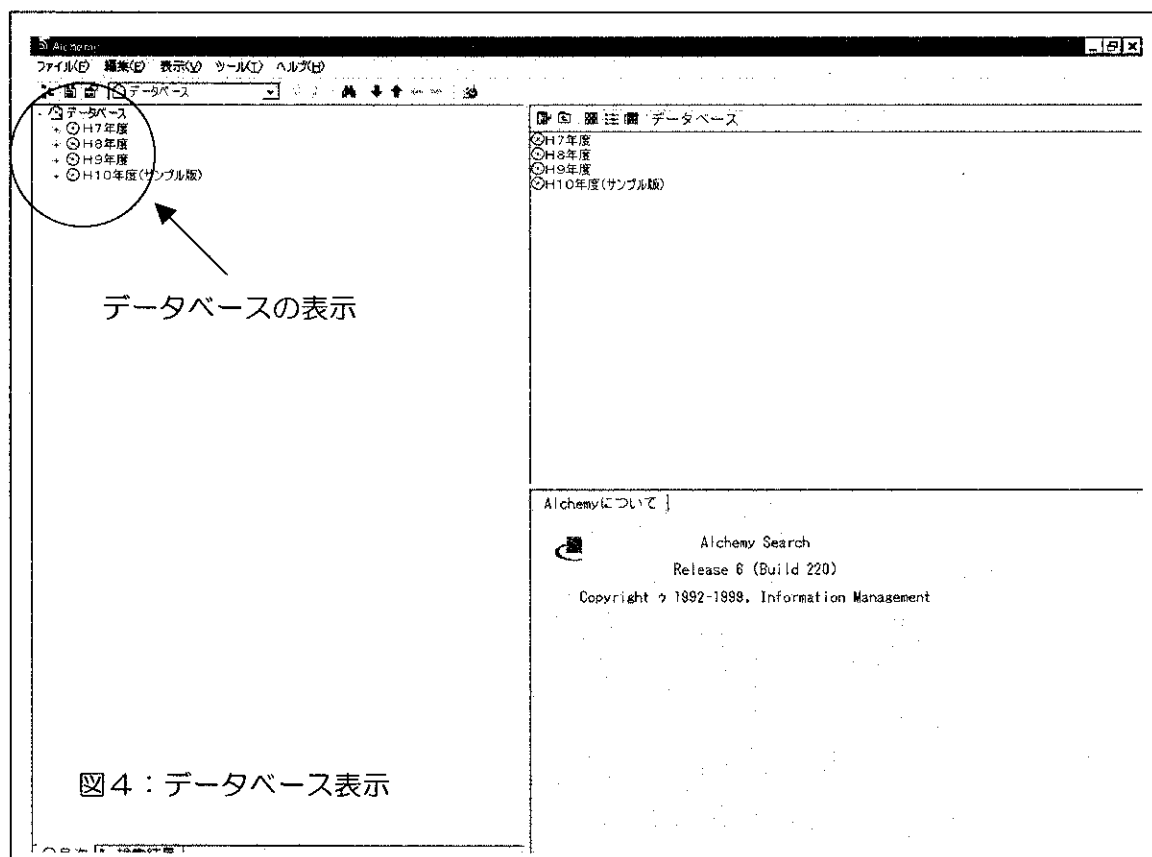


図 3：CD 版のファイル構成

2) データベースの表示

この画面は、今年度作成した過去分心身障害研究報告書の平成 7～9 年度版と平成 10 年

度子ども家庭総合研究報告書のサンプル版をDVD-RAMにコピーしたものである。したがって、実際にリリースするCD版（各年度1枚）と若干異なることをお断りしておく。



データベースを読み込むと、図4の画面になる。この画面では、平成7年度、平成8年度、平成9年度、平成10年度（サンプル版）データベースが読み込まれていることを表している。

3) 各年度の主任研究班と階層構造 (図5、6)

各データベースの目次は、各年度の主任研究班の研究テーマ、分担研究者の研究テーマ、研究協力者の研究テーマごとに、階層構造になっており、主任研究班の下の階層に分担研究班があり、その下の階層に研究協力者が表示される。

4) 次に、検索機能を検証してみる。

各データベースの目次（主任研究班）が表示された状態で、メニューバーの検索メニューまたは検索アイコンを開いてみると、図6のようになる。

ドキュメント情報とドキュメント内容の両者を

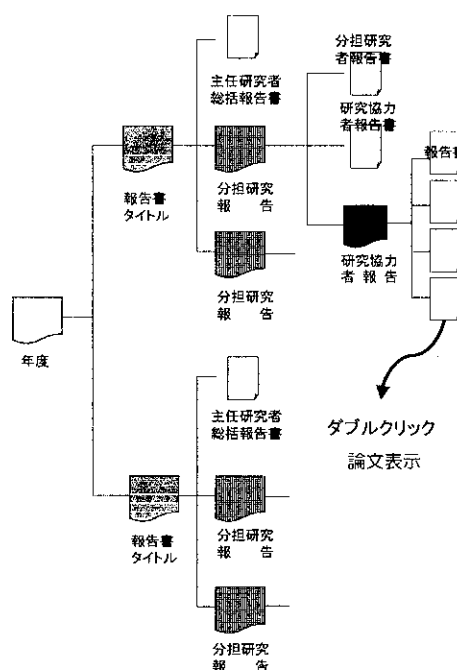
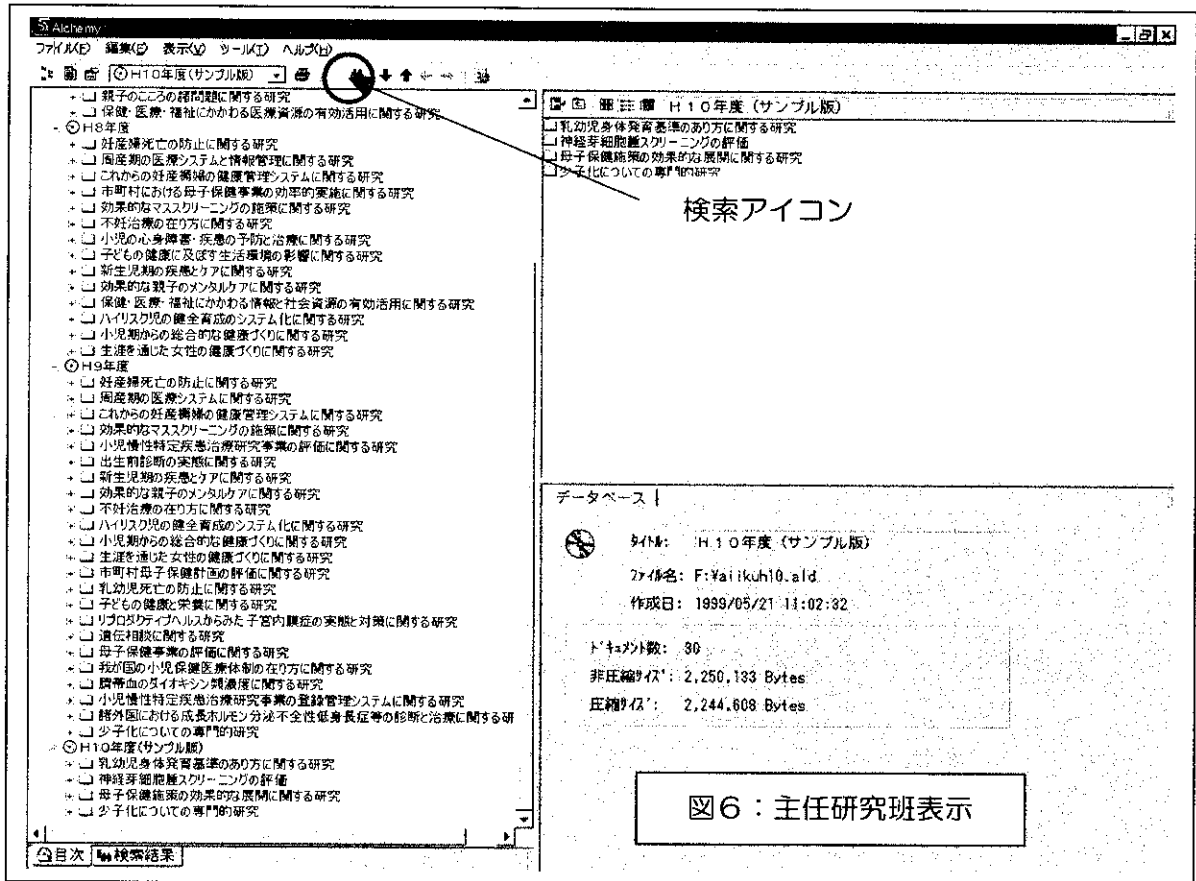


図5：データベース階層構造

検索の対象とするように、チェックボタンにチェックを入れておく。ドキュメント情報フィールドとは、各研究タイトル、著者、キーワードを意味しており、ドキュメント内容とは、要約、目的、まとめなどOCRで切り出しテキスト化して、論文の最終ページにリンクしてあるテキスト文を意味する。また、新規分では、論文の7~8割は全文テキストで取り込み、PDFファイル化してあるので、全文の文字列検索が可能である。



次に、図7に示したように、検索したい文字列を検索式のフィールドに入力する。検索にはアスタリスク(*)やワルドカード(?)が使用できる。即ち、AAA*は先頭一致を意味する。全ての文中からAAAという文字列を検索したいのなら、*AAA*を用いる。また、ANDとORの演算子を用いることもできる。

例題として、*子育て* and *支援* で検索してみる。文中に「子育て」と「支援」という2つの単語がある報告書を検索することを意味している。

結果は、図8に示した。各データベースごとに抽出されたドキュメントの数と研究報告書のテーマを表示している。

5) 報告書内容の表示と目次との関連づけ

検索されたドキュメントの先頭のものが高ライトされて、右側の上段のウィンドウに報告書の内容がPDFファイルで表示される。

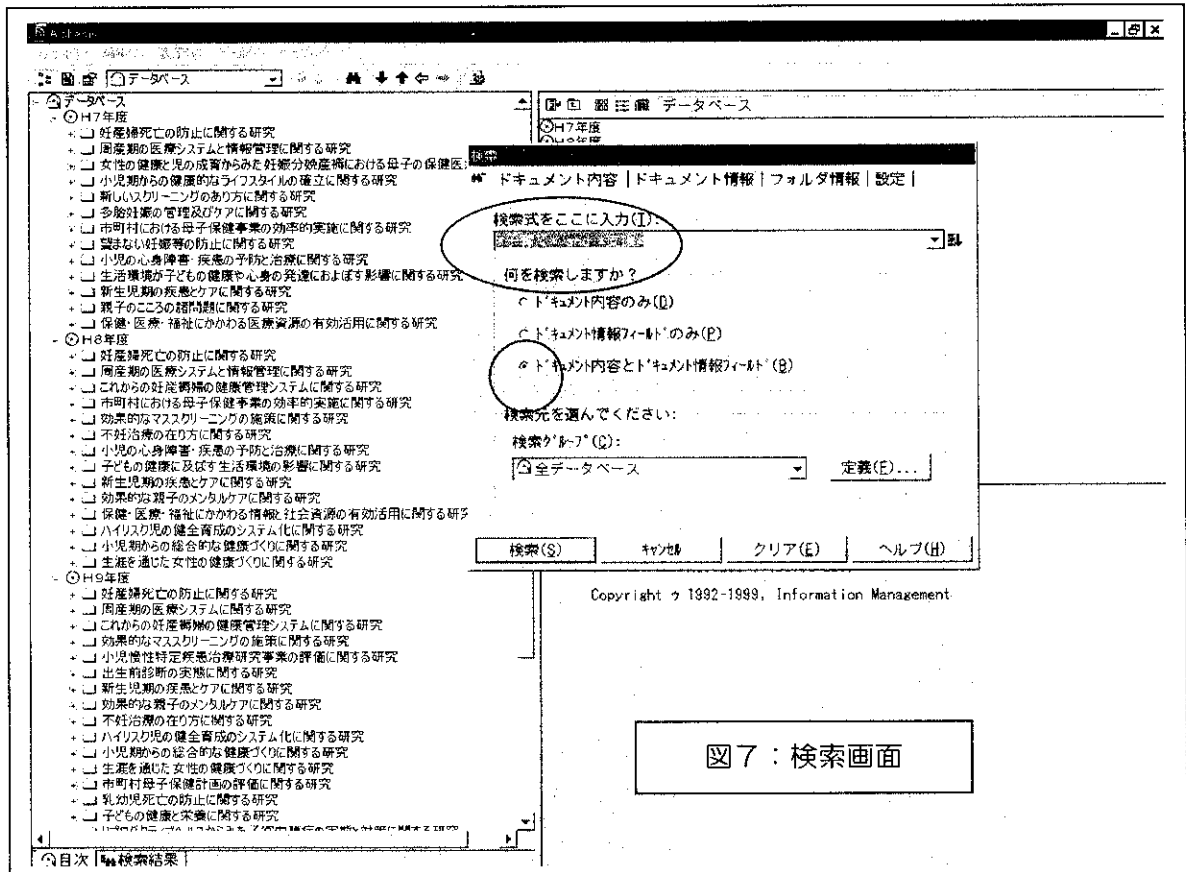
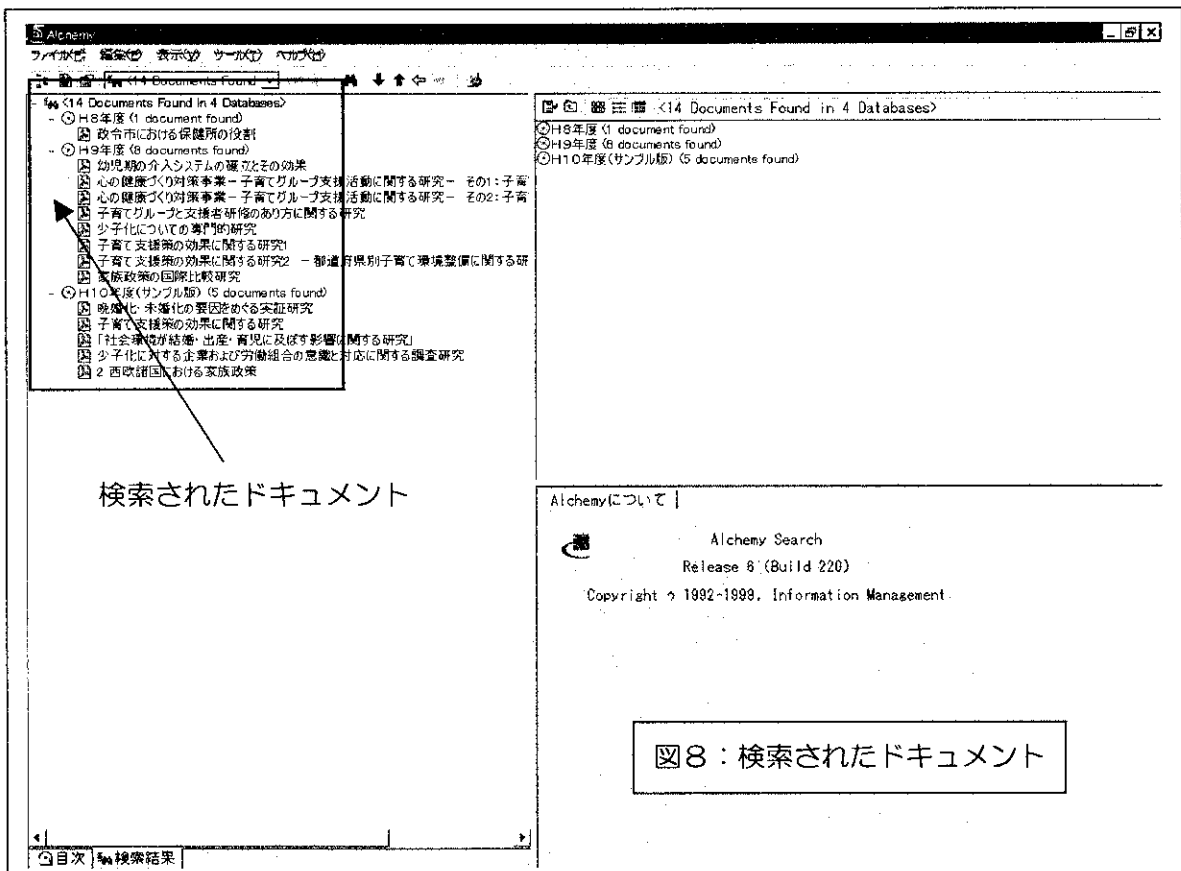
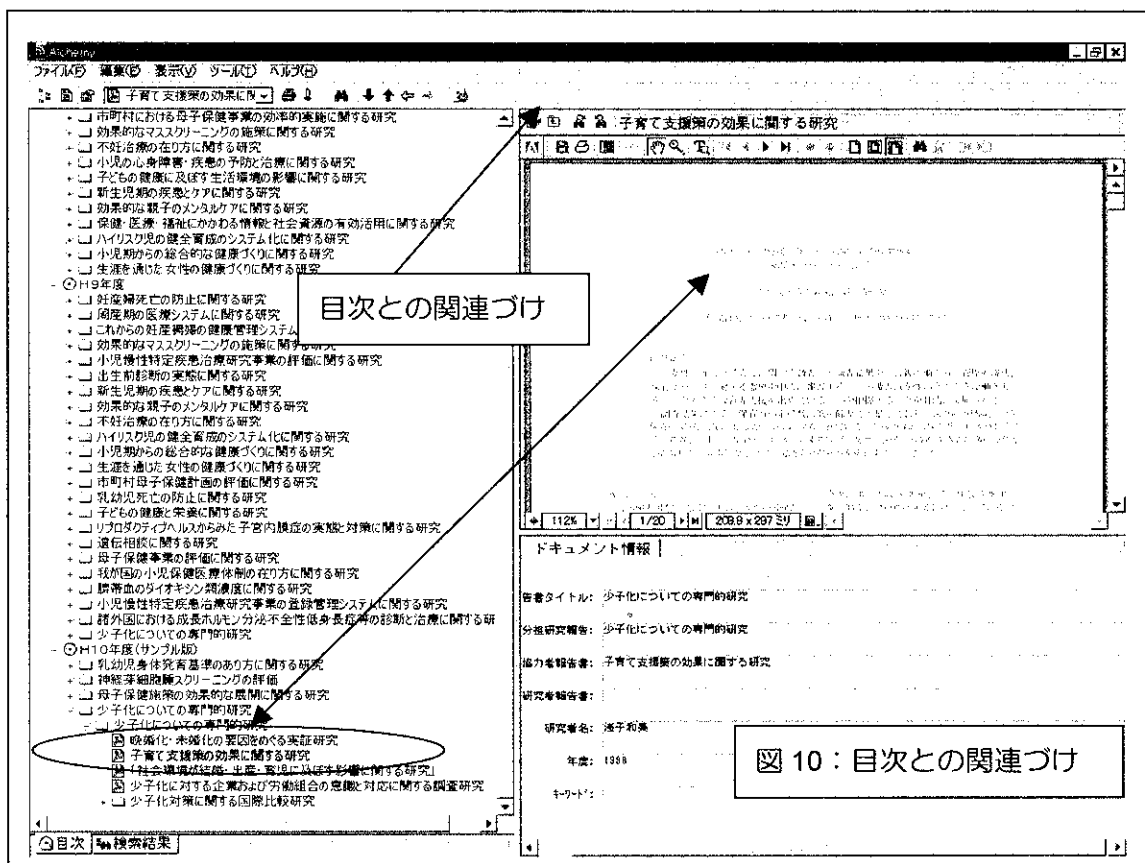
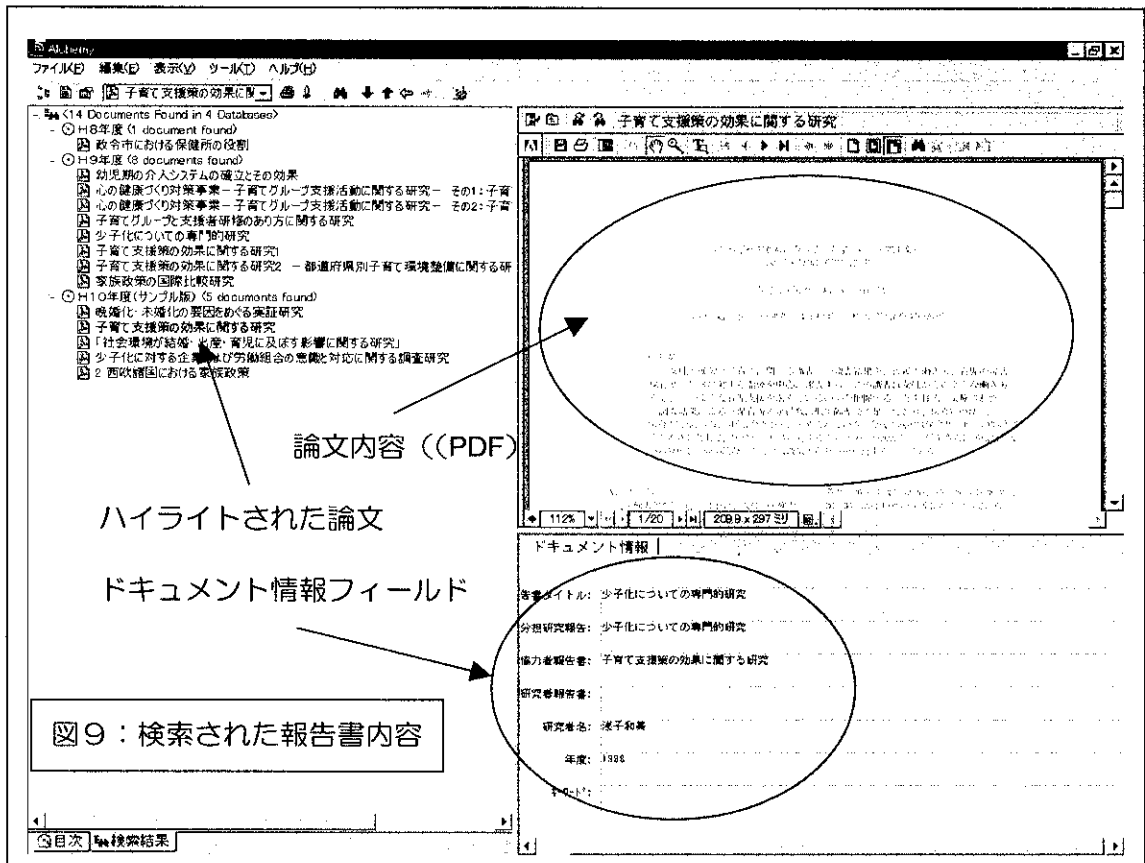


図7：検索画面



検索されたドキュメント

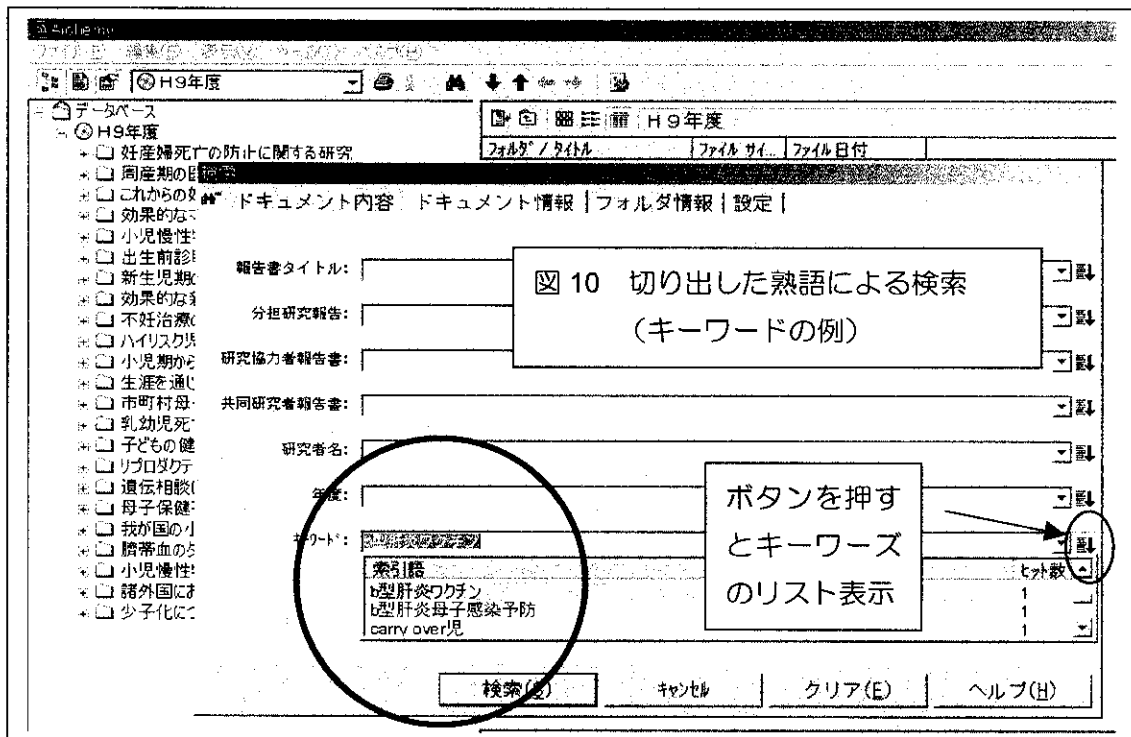
図8：検索されたドキュメント



目次との関連づけアイコンをクリックすると、PDF ファイルで表示されている論文が報告書のどの階層に属しているかがわかる。つまり、研究班のどの位置に属しているものかわかる。

6) もう一つの検索機能

ドキュメントフィールドを用いる方法がある。この方法では、任意の文字列を入力してもよいが、各ドキュメントフィールドの表示ボタンを押すと切り出された文字列が表示される。この表示された文字列から検索したい文字列を選択して検索ができる。

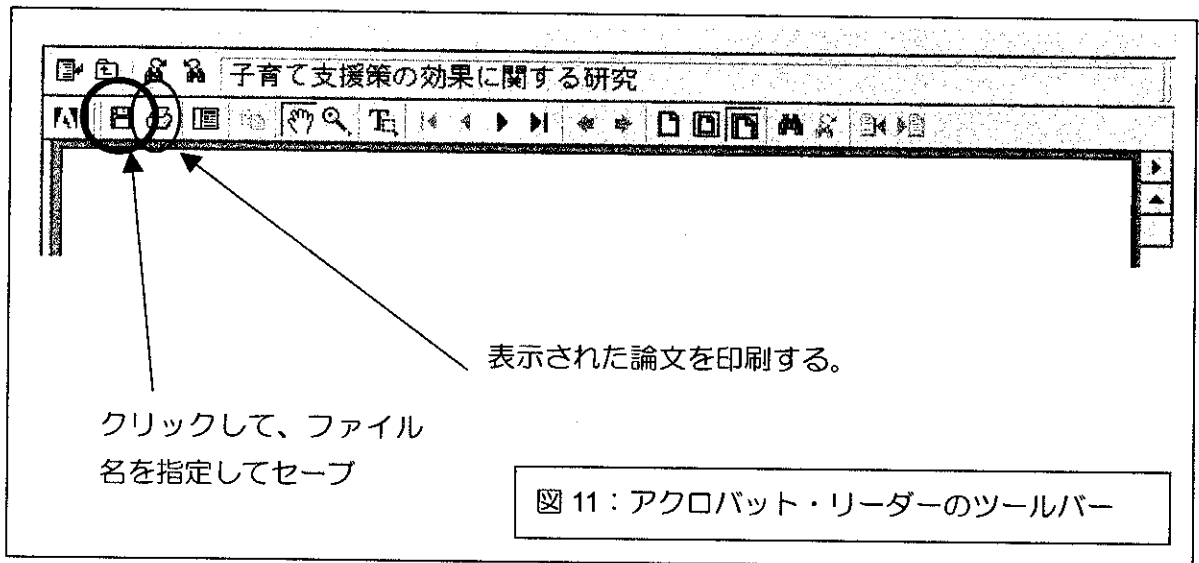


PDF ファイルで表示されている論文は、アクロバットの「ファイルアイコン」をクリックし、ファイルとしてセーブすることができる。もう一つの方法は、Alchemy のファイルメニューから、「ファイルの取り出し」を選択して、ファイルをセーブするドライブとディレクトリーを指定すれば、PDF ファイルとしてセーブすることができる。

7) 表示された論文の印刷は、アクロバット・リーダーの印刷アイコンをクリックすることで印刷が可能である。

8) 最後に、印刷された PDF ファイルのサンプルを添付しておく（巻末資料）。

このサンプルは報告書を画像ファイルとして取り込んだ過去分データのサンプルである。報告書の最終ページに OCR で取り込み、検索用の文字列として添付したテキスト文章が付属している。この部分は、OCR ソフトにより、機械的に処理しているので、誤字や脱字のみられることがあり、取り扱いには注意が必要である。



D. 考察

過去の「心身障害研究」および今年度から実施される厚生科学研究「子ども家庭総合研究」の成果が広く有効に活用されるためには、より多くの関係者に報告書が配布され、研究内容を手軽に入手できることが重要である。

従来の厚生省心身障害研究報告書は優れた研究成果が掲載されているにも関わらず、限られた行政機関や研究機関に配布されたのみで、多くの学徒、研究者、実地で活躍する医療・保健・福祉の専門家の目に触れにくいという難点を抱えていた。

そこで、これらの報告書の普及と有効活用を図るためには、報告書を電子化し、検索機能を備えたデータベースを構築するのが最も有効な方法であると考えられた。電子化された報告書は、従来の印刷物に比べ、コンパクトな媒体に大量のデータを収納することが可能であり、いったん、電子化すれば需要に合わせて、その後の複製も容易であり、保管場所などの制約と増刷に多くの費用を要する印刷物に比べれば、対費用効果も優れている。また、データベース化により、大量の論文の中から、必要な資料を検索することができるなど、印刷物で保管するより遙かに利便性が高いと考えられる。

今年度は、昨年度検討し確立した電子化の方法論にしたがって、実際のデータベース化作業に着手した。まだ、若干、技術的問題を残してはいるが、十分実用に供することができるデータベースに仕上がっている。

昨年度研究の要約を加えながら今年度研究の成果と問題点について述べると次のようになる。

1) 電子データ化の方法論

報告書の電子データ化にあたって、既に製本されている過去分「心身障害研究報告書」とこれからの「子ども家庭総合研究報告書」に分けて検討した。

過去分の心身障害研究報告書については、電子化作業の効率と経済性から、画像データを使用したデータベース作成が妥当であると考えられた。方法論は報告書を解体し論文の各ページをイ

メーグスキャナーで読みとり、画像ファイルとして保管する方法を用いた。

新規の「子ども家庭総合研究報告書」は、7から8割は電子データとして収集が可能であると考えられる。しかし、そのファイル形式は多様な機種やソフトで作成されていることは過去の研究でも明白であり¹⁾、その多様な電子データを統一したファイル形式に変換することが必要になる。また、電子データで提出されない報告書、即ち印刷物のみの場合や変換不可能なワープロ専用機などで作成された報告書については、過去分報告書の電子化と同様に、画像データとして収録する方法を併用する必要がある。また、指定ソフトで作成されていない図表なども画像データとして収録し、当該論文ファイルとリンクさせる必要がある。

電子化した大量データは、収納するファイル形式を選択する必要がある。多くの情報をコンパクトに収納し、利便性を高めるためには、ファイルサイズが小さいことも条件であり、電子データのファイル形式は、現時点ではPDFファイルが最適と考えられる。これはファイルサイズが小さく、印刷イメージと近似した形態で収録でき、どんなプリンターでもオリジナルどおり印刷でき、テキスト部分・イメージ部分ともに電子的切り貼りが可能である。また、データベース化を図るとき、PDFファイルに収録したテキスト部分は、文字列として検索が可能であり、PDFファイルに対応した検索ソフトを使用すれば、ほぼ全文検索の機能を備えることができる。

また、PDFを表示・印刷するビューワーソフト（Acrobat Reader）は、各種のプラットフォームで閲覧することが可能であり、イメージファイルもPDFに変換することによりファイルサイズが小さくなり、大量データの収録が可能になる。

2) 検索用ソフトの選定

データベースを構築する上で、ポイントになるのはオーサリングと検索機能であり、昨年度の検討結果を踏まえ、Alchemy release 6を用いて実用化に着手した。

Alchemyは、現在さらに機能アップし、ODBCドライバーに対応し、Visual Basicなどを用いたカスタマイズ機能が充実したVersion 6がリリースされ、日本語バージョンも完成された。しかしながら、問題もあり、わが国で医療関係機関を中心に普及しているMacintosh OSには対応しておらず、今後の検討課題として残っている。

3) 今年度の研究成果と問題点

今回、Alchemy Release 6日本語バージョンにより、PDFファイル化した報告書を用いてデータベースを作成した。過去分心身障害研究報告書は昭和50年～平成9年度までの23年間についてデータベース化を行い、平成10年度新規分子も家庭総合研究報告書のデータベース化に着手しており、ほぼ完成の域に達している。

研究報告書のデータベース化に当たって、もっとも重要な機能は、自由語での高い検索機能である。さらに、報告書という印刷物の電子化であり、電子化された書物という体裁も必要になる。

今回作成したAlchemyによるデータベースは、報告書の各論文を、主任研究班、分担研究班、研究協力者、さらに共同研究者などの階層構造に分けて、比較的自由に設定されており、報告書の目次のイメージを再現できている点で、利用者にとって原本の報告書と変わらぬ自然な構造で構築できている。検索機能には、多少問題はあるが、検索文字列を演算子で組み合わせて検索す