

2026年版DRAFT

NDB・DPC等データベース研究人材育成Webinar

SQLを用いたレセプトデータのハンドリング
上級編

東京大学大学院医学系研究科臨床疫学・経済学
松居 宏樹

ここ最近の研究成果

1. Ishimaru M, Matsui H, Ono S, Hagiwara Y, Morita K, Yasunaga H. Preoperative oral care and effect on postoperative complications after major cancer surgery. *British Journal of Surgery*. 2018 Oct 12;105(10):1685-96.

2. Hashimoto Y, Matsui H, Michihata N, Ishimaru M, Yasunaga H, Aihara M, et al. Incidence of sympathetic ophthalmia after incising events: a national database study in Japan. *Ophthalmology*. 2021 Sep 21;50161-6420(21)00719-3.

3. Takeuchi Y, Kumamaru H, Hagiwara Y, Matsui H, Yasunaga H, Miyata H, et al. Sodium-glucose cotransporter-2 inhibitors and the risk of urinary tract infection among diabetic patients in Japan: Target trial emulation using a nationwide administrative claims database. *Diabetes Obes Metab*. 2021 Jun;23(6):1379-88.

4. Ishimaru M, Ono S, Morita K, Matsui H, Hagiwara Y, Yasunaga H. Prevalence, Incidence Rate, and Risk Factors of Medication-Related Osteonecrosis of the Jaw in Patients With Osteoporosis and Cancer: A Nationwide Population-Based Study in Japan. *Journal of Oral and Maxillofacial Surgery*. 2022;80(4):714-27.

5. Kasajima M, Eggleston K, Kusaka S, Matsui H, Tanaka T, Son BK, et al. Projecting prevalence of frailty and dementia and the economic cost of care in Japan from 2016 to 2043: a microsimulation modelling study. *The Lancet Public Health*. 2022;7(5):e558-68.

6. Kimura Y, Suzuki M, Ito T, Hashimoto Y, Kumazawa R, Ishimaru M, et al. Epidemiology of severe childhood asthma in Japan: A nationwide descriptive study. *Allergy*. 2024 Jan 7;all.16008.

7. Kimura Y, Jo T, Hashimoto Y, Kumazawa R, Ishimaru M, Matsui H, et al. Epidemiology of patients with lymphangioleiomyomatosis: A descriptive study using the national database of health insurance claims and specific health checkups of Japan. *Respir Investig*. 2024 May;62(3):494-502.



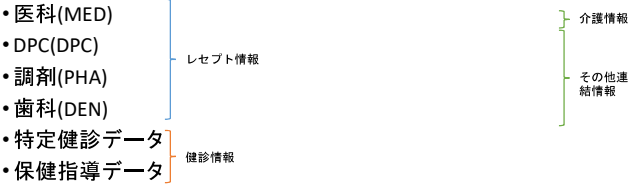
この授業では

- 大規模なレセプトデータベースを解析する場合に当たって、
- レセプトデータベースの構造を理解する。
 - 解析用データセットの設計・構築の流れを学習する。



NDBデータ構造
NDBデータの大まかな構造

- レセプト部分と健診情報部分がある。
- また、介護情報等との接続も可能になった。



今回の演習では主として医療レセプトの
データハンドリングを学ぶ



学習サイト

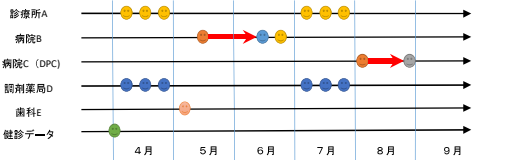
- 学習を円滑に進めるため、一時的なWebサイトを公開。

準備中



NDBデータ構造
NDBデータの大まかな構造

患者のエピソードで考えると以下の様になる。



- 外来受診 ● 調剤薬局 ● 歯科受診
- 入院開始 ● 生存退院 ● 死亡退院
- 健診受診

4月：診療所Aに外来受診3回、同日調剤、健診受診
5月：歯科E受診、病院B入院
6月：病院B退院、同病院にて外来フォローアップ
7月：診療所Aに外来受診3回、同日調剤
8月：病院CにDPC入院、同月死亡退院

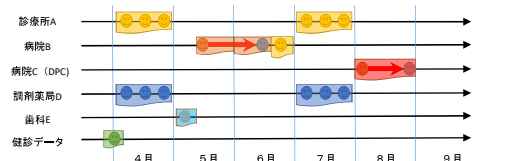


はじめに
NDB とは何か

- 正式名称：レセプト情報・特定健診等情報 データベース
- 利用目的は全国医療費適正化計画及び都道府県医療費適正化計画の作成、実施及び評価に資するため（高齢者の医療の確保に関する法律 第16条）
- 保有：厚生労働大臣
- 収集元：審査支払機関
- 内容
 - レセプトデータ 年間約18億7000万件（H28年度）
 - 特定健診・保健指導データ 年間約2,730万件（H28年度）
 - Administrative Claims Database の一種
- 窓口など：
https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/iryohoken/resepto/index.html



NDBデータ構造
NDBデータの大まかな構造



- 資料レセ（外来） 医科レセ（入院）
- DPCレセ 調剤レセ
- 歯科レセ 健診データ

4月：診療所Aから医科レセ（外来）×1、
調剤薬局Dから調剤レセ×1、健診データ×1
5月：歯科Eから歯科レセ×1、病院Bから医科レセ（入院）×1
6月：病院Bから医科レセ（入院）×1、医科レセ（外来）×1
7月：診療所Aから医科レセ（外来）×1、調剤薬局Dから調剤レセ×1
8月：病院BからDPCレセ×1



レセプトデータの構造化

- データを検索してみる
- CSV形式のレセプトをエクセルで開く

"Jmedsample/0_COMMON0001_MED/11_RECORDINFO_MED.CSV"

- サンプルデータをレセプト単位に切り分ける

MNレコードが切れ目

- サンプルデータから、医科レセの医薬品処方履歴を集める。
- "サンプル" → "さんの"尿素クリーム10%「フジナガ」"の処方日は？



東京大学
THE UNIVERSITY OF TOKYO

- # レセプトデータの構造化
- データを検索してみる
 - CSV形式のレセプトをエクセルで開く
- "Jmedsample/0_COMMON0001_MED/11_RECORDINFO_MED.CSV"
- サンプルデータをレセプト単位に切り分ける
- ## MNレコードが切れ目
- サンプルデータから、医科レセの医薬品処方履歴を集める。
 - "サンプル ー"さんの"尿素クリーム10%「フジナガ」"の処方日は？
- 
- 東京大学
THE UNIVERSITY OF TOKYO



特別抽出仕様NDB データベース

postgresql 上にサンプルレセプトを基にレセプトデータベースを作成

- 特別抽出のデータ仕様に合わせたデータベースを構築
 - ・ テーブル名・カラム名は先述の仕様を基に決定
 - ・ サンプルレセプトをデータベースに格納
- データベースソフトウェアはpostgresql ver.14.6 を用いた。
 - ・ オンサイトセンターでの採用
 - ・ Redshift などの種々のクラウドマネージドデータベースと構文が似ている。
 - ・ オープンソースソフトウェアで導入しやすい。

The diagram illustrates a cloud-based architecture. A central PostgreSQL database instance is shown with connections to Amazon S3 buckets (labeled 'sample-receipt' and 'receipt-report'), an IAM role, and an AWS Lambda function. The database also connects to an Amazon EC2 instance via a VPN connection. Other components include an Amazon RDS instance, an Amazon ElastiCache Redis instance, and an Amazon CloudWatch console.

特別抽出仕様NDB データベース

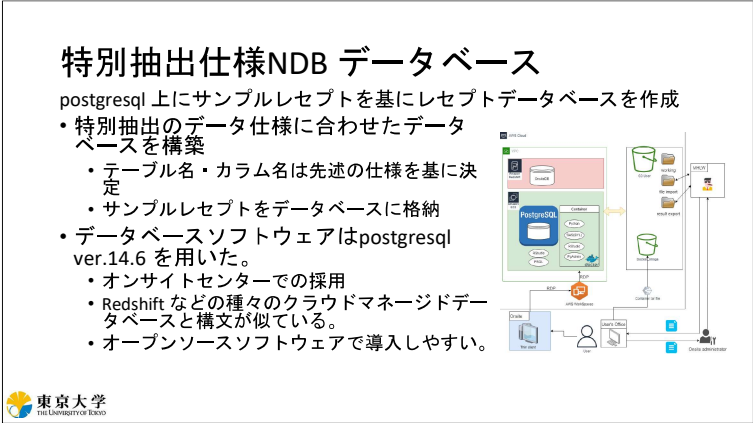
postgresql 上にサンプルレセプトを基にレセプトデータベースを作成

- 特別抽出のデータ仕様に合わせたデータベースを構築
 - ・ テーブル名・カラム名は先述の仕様を基に決定
 - ・ サンプルレセプトをデータベースに格納
- データベースソフトウェアはpostgresql ver.14.6 を用いた。
 - ・ オンサイトセンターでの採用
 - ・ Redshift などの種々のクラウドマネージドデータベースと構文が似ている。
 - ・ オープンソースソフトウェアで導入しやすい。

The diagram illustrates the database architecture. It shows PostgreSQL connected via JDBC to a Java application. The Java app interacts with a 'sample-receipt' database containing tables like 'tbl_receipt_header' and 'tbl_receipt_detail'. Data flows from the sample database to a 'receipt-database' which also contains similar tables. This database is linked to a 'Data Lake' and a 'Data Warehouse' (Snowflake). A user icon is shown interacting with the system through a web browser.

東京大学
THE UNIVERSITY OF TOKYO

- [illegible]



レセプトデータの構造化

- レセプトデータは生のままでは扱いにくい。
- 行ごとに書いてある情報が違う
- 行の種類ごとに分けてデータベース化しよう。

- # レセプトデータの構造化
- レセプトデータは生のままでは扱いにくい。
 - 行ごとに書いてある情報が違う
 - 行の種類ごとに分けてデータベース化しよう。



NDBデータの大まかな構造

- 複数のテーブル(右図)で1レセの情報がとなる。
- 匿名化個人IDは各REに含まれる、匿名化ID1と匿名化ID2の2つ
- 他のレセプト(MED, DPC, PHA, DEN)情報(や健診情報)とはレセプトIDで接続可能
- 月単位で登録されるため、月またぎ入院や月内外来受診の処理が煩雑

レセプト共通 (RE)

匿名化個人ID (N1)

匿名化個人ID (N2)

匿名化個人ID (N3)

匿名化個人ID (N4)

※DPCレセはREが複数値

NDBデータの大まかな構造

- 複数のテーブル(右図)で1レセの情報がとなる。
- 匿名化個人IDは各REに含まれる、匿名化ID1と匿名化ID2の2つ
- 他のレセプト(MED, DPC, PHA, DEN)情報(や健診情報)とはレセプトIDで接続可能
- 月単位で登録されるため、月またぎ入院や月内外来受診の処理が煩雑

レセプト共通 (RE)

匿名化個人ID (N1)

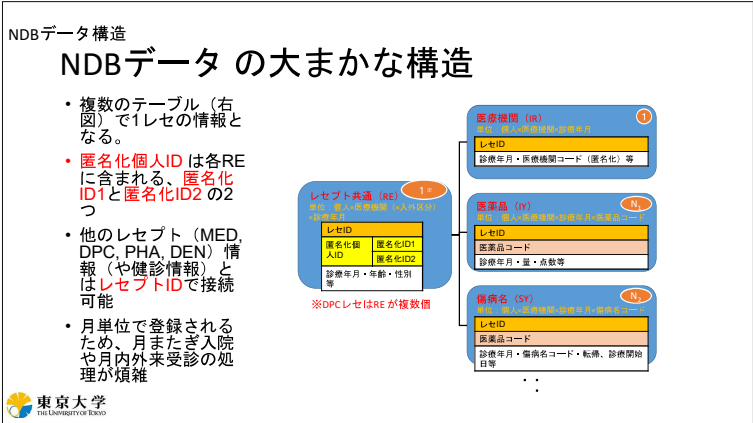
匿名化個人ID (N2)

匿名化個人ID (N3)

匿名化個人ID (N4)

※DPCレセはREが複数値

- # NDBデータの大まかな構造
- 複数のテーブル(右図)で1レセの情報がとなる。
 - 匿名化個人IDは各REに含まれる、匿名化ID1と匿名化ID2の2つ
 - 他のレセプト(MED, DPC, PHA, DEN)情報(や健診情報)とはレセプトIDで接続可能
 - 月単位で登録されるため、月またぎ入院や月内外来受診の処理が煩雑
-
- レセプト共通 (RE)
- 匿名化個人ID (N1)
- 匿名化個人ID (N2)
- 匿名化個人ID (N3)
- 匿名化個人ID (N4)
- ※DPCレセはREが複数値



実際の研究とデータ構造

- 構造化されたデータの研究利用
NDBのテーブル構造はあまり大っぴらにされていない。
[NDBの提供について（厚労省）](#)のサイト内にある、抽出依頼テンプレートに詳細がある。

- 抽出依頼テンプレート：当方と相談の上、必要に応じて作成してください
 - ▶ [X](#) 中出依頼テンプレート(注意) | DX 58形式:1,210601 | [①](#)
 - ▶ [X](#) 中出依頼テンプレート(諸君) | DX 58形式:1,720801 | [②](#)
 - ▶ [X](#) 中出依頼テンプレート(集計) | DX 58形式:1,530601 | [③](#)

 東京大学
the University of Tokyo

- # 実際の研究とデータ構造
- 構造化されたデータの研究利用
NDBのテーブル構造はあまり大っぴらにされていない。
[NDBの提供について（厚労省）](#)のサイト内にある、抽出依頼テンプレートに詳細がある。
- 抽出依頼テンプレート：当方と相談の上、必要に応じて作成してください
 - ▶ [X](#) 中出依頼テンプレート(注意) [XLS形式:1,210KB] 
 - ▶ [X](#) 中出依頼テンプレート(簡易) [XLS形式:1,720KB] 
 - ▶ [X](#) 中出依頼テンプレート(集計) [XLS形式:1,530KB] 
- 

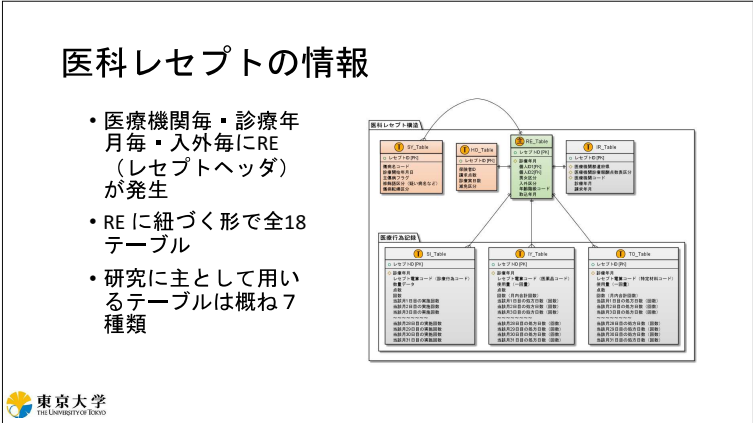
東京大学
the University of Tokyo

- # 実際の研究とデータ構造
- 構造化されたデータの研究利用
NDBのテーブル構造はあまり大っぴらにされていない。
[NDBの提供について（厚労省）](#)のサイト内にある、抽出依頼テンプレートに詳細がある。
- 抽出依頼テンプレート：当方と相談の上、必要に応じて作成してください
 - ▶ [X](#) 中出依頼テンプレート(注意) | DX 58形式:1,210601 | [①](#)
 - ▶ [X](#) 中出依頼テンプレート(諸君) | DX 58形式:1,720801 | [②](#)
 - ▶ [X](#) 中出依頼テンプレート(集計) | DX 58形式:1,530601 | [③](#)
- 

東京大学
the University of Tokyo

[illegible]

- [illegible]

[illegible]

- [illegible]

DB仕様は機械判読可能な形で

- エクセル内に提供ファイルのコラム型情報あり

項目	説明	単位	科目	履修科目	履修科目	履修科目
1	基礎知識 (1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15) (16) (17) (18) (19) (20) (21) (22) (23) (24) (25) (26) (27) (28) (29) (30) (31) (32) (33) (34) (35) (36) (37) (38) (39) (40) (41) (42) (43) (44) (45) (46) (47) (48) (49) (50) (51) (52) (53) (54) (55) (56) (57) (58) (59) (60) (61) (62) (63) (64) (65) (66) (67) (68) (69) (70) (71) (72) (73) (74) (75) (76) (77) (78) (79) (80) (81) (82) (83) (84) (85) (86) (87) (88) (89) (90) (91) (92) (93) (94) (95) (96) (97) (98) (99) (100) (101) (102) (103) (104) (105) (106) (107) (108) (109) (110) (111) (112) (113) (114) (115) (116) (117) (118) (119) (120) (121) (122) (123) (124) (125) (126) (127) (128) (129) (130) (131) (132) (133) (134) (135) (136) (137) (138) (139) (140) (141) (142) (143) (144) (145) (146) (147) (148) (149) (150) (151) (152) (153) (154) (155) (156) (157) (158) (159) (160) (161) (162) (163) (164) (165) (166) (167) (168) (169) (170) (171) (172) (173) (174) (175) (176) (177) (178) (179) (180) (181) (182) (183) (184) (185) (186) (187) (188) (189) (190) (191) (192) (193) (194) (195) (196) (197) (198) (199) (200) (201) (202) (203) (204) (205) (206) (207) (208) (209) (210) (211) (212) (213) (214) (215) (216) (217) (218) (219) (220) (221) (222) (223) (224) (225) (226) (227) (228) (229) (230) (231) (232) (233) (234) (235) (236) (237) (238) (239) (240) (241) (242) (243) (244) (245) (246) (247) (248) (249) (250) (251) (252) (253) (254) (255) (256) (257) (258) (259) (260) (261) (262) (263) (264) (265) (266) (267) (268) (269) (270) (271) (272) (273) (274) (275) (276) (277) (278) (279) (280) (281) (282) (283) (284) (285) (286) (287) (288) (289) (290) (291) (292) (293) (294) (295) (296) (297) (298) (299) (300) (301) (302) (303) (304) (305) (306) (307) (308) (309) (310) (311) (312) (313) (314) (315) (316) (317) (318) (319) (320) (321) (322) (323) (324) (325) (326) (327) (328) (329) (330) (331) (332) (333) (334) (335) (336) (337) (338) (339) (340) (341) (342) (343) (344) (345) (346) (347) (348) (349) (350) (351) (352) (353) (354) (355) (356) (357) (358) (359) (360) (361) (362) (363) (364) (365) (366) (367) (368) (369) (370) (371) (372) (373) (374) (375) (376) (377) (378) (379) (380) (381) (382) (383) (384) (385) (386) (387) (388) (389) (390) (391) (392) (393) (394) (395) (396) (397) (398) (399) (400) (401) (402) (403) (404) (405) (406) (407) (408) (409) (410) (411) (412) (413) (414) (415) (416) (417) (418) (419) (420) (421) (422) (423) (424) (425) (426) (427) (428) (429) (430) (431) (432) (433) (434) (435) (436) (437) (438) (439) (440) (441) (442) (443) (444) (445) (446) (447) (448) (449) (450) (451) (452) (453) (454) (455) (456) (457) (458) (459) (460) (461) (462) (463) (464) (465) (466) (467) (468) (469) (470) (471) (472) (473) (474) (475) (476) (477) (478) (479) (480) (481) (482) (483) (484) (485) (486) (487) (488) (489) (490) (491) (492) (493) (494) (495) (496) (497) (498) (499) (500) (501) (502) (503) (504) (505) (506) (507) (508) (509) (510) (511) (512) (513) (514) (515) (516) (517) (518) (519) (520) (521) (522) (523) (524) (525) (526) (527) (528) (529) (530) (531) (532) (533) (534) (535) (536) (537) (538) (539) (540) (541) (542) (543) (544) (545) (546) (547) (548) (549) (550) (551) (552) (553) (554) (555) (556) (557) (558) (559) (560) (561) (562) (563) (564) (565) (566) (567) (568) (569) (570) (571) (572) (573) (574) (575) (576) (577) (578) (579) (580) (581) (582) (583) (584) (585) (586) (587) (588) (589) (590) (591) (592) (593) (594) (595) (596) (597) (598) (599) (600) (601) (602) (603) (604) (605) (606) (607) (608) (609) (610) (611) (612) (613) (614) (615) (616) (617) (618) (619) (620) (621) (622) (623) (624) (625) (626) (627) (628) (629) (630) (631) (632) (633) (634) (635) (636) (637) (638) (639) (640) (641) (642) (643) (644) (645) (646) (647) (648) (649) (650) (651) (652) (653) (654) (655) (656) (657) (658) (659) (660) (661) (662) (663) (664) (665) (666) (667) (668) (669) (670) (671) (672) (673) (674) (675) (676) (677) (678) (679) (680) (681) (682) (683) (684) (685) (686) (687) (688) (689) (690) (691) (692) (693) (694) (695) (696) (697) (698) (699) (700) (701) (702) (703) (704) (705) (706) (707) (708) (709) (710) (711) (712) (713) (714) (715) (716) (717) (718) (719) (720) (721) (722) (723) (724) (725) (726) (727) (728) (729) (730) (731) (732) (733) (734) (735) (736) (737) (738) (739) (740) (741) (742) (743) (744) (745) (746) (747) (748) (749) (750) (751) (752) (753) (754) (755) (756) (757) (758) (759) (760) (761) (762) (763) (764) (765) (766) (767) (768) (769) (770) (771) (772) (773) (774) (775) (776) (777) (778) (779) (780) (781) (782) (783) (784) (785) (786) (787) (788) (789) (790) (791) (792) (793) (794) (795) (796) (797) (798) (799) (800) (801) (802) (803) (804) (805) (806) (807) (808) (809) (810) (811) (812) (813) (814) (815) (816) (817) (818) (819) (820) (821) (822) (823) (824) (825) (826) (827) (828) (829) (830) (831) (832) (833) (834) (835) (836) (837) (838) (839) (840) (841) (842) (843) (844) (845) (846) (847) (848) (849) (850) (851) (852) (853) (854) (855) (856) (857) (858) (859) (860) (861) (862) (863) (864) (865) (866) (867) (868) (869) (870) (871) (872) (873) (874) (875) (876) (877) (878) (879) (880) (881) (882) (883) (884) (885) (886) (887) (888) (889) (890) (891) (892) (893) (894) (895) (896) (897) (898) (899) (900) (901) (902) (903) (904) (905) (906) (907) (908) (909) (910) (911) (912) (913) (914) (915) (916) (917) (918) (919) (920) (921) (922) (923) (924) (925) (926) (927) (928) (929) (930) (931) (932) (933) (934) (935) (936) (937) (938) (939) (940) (941) (942) (943) (944) (945) (946) (947) (948) (949) (950) (951) (952) (953) (954) (955) (956) (957) (958) (959) (960) (961) (962) (963) (964) (965) (966) (967) (968) (969) (970) (971) (972) (973) (974) (975) (976) (977) (978) (979) (980) (981) (982) (983) (984) (985) (986) (987) (988) (989) (990) (991) (992) (993) (994) (995) (996) (997) (998) (999) (1000) (1001) (1002) (1003) (1004) (1005) (1006) (1007) (1008) (1009) (1010) (1011) (1012) (1013) (1014) (1015) (1016) (1017) (1018) (1019) (1020) (1021) (1022) (1023) (1024) (1025) (1026) (1027) (1028) (1029) (1030) (1031) (1032) (1033) (1034) (1035) (1036) (1037) (1038) (1039) (1040) (1041) (1042) (1043) (1044) (1045) (1046) (1047) (1048) (1049) (1050) (1051) (1052) (1053) (1054) (1055) (1056) (1057) (1058) (1059) (1060) (1061) (1062) (1063) (1064) (1065) (1066) (1067) (1068) (1069) (1070) (1071) (1072) (1073) (1074) (1075) (1076) (1077) (1078) (1079) (1080) (1081) (1082) (1083) (1084) (1085) (1086) (1087) (1088) (1089) (1090) (1091) (1092) (1093) (1094) (1095) (1096) (1097) (1098) (1099) (1100) (1101) (1102) (1103) (1104) (1105) (1106) (1107) (1108) (1109) (1110) (1111) (1112) (1113) (1114) (1115) (1116) (1117) (1118) (1119) (1120) (1121) (1122) (1123) (1124) (1125) (1126) (1127) (1128) (1129) (1130) (1131) (1132) (1133) (1134) (1135) (1136) (1137) (1138) (1139) (1140) (1141) (1142) (1143) (1144) (1145) (1146) (1147) (1148) (1149) (1150) (1151) (1152) (1153) (1154) (1155) (1156) (1157) (1158) (1159) (1160) (1161) (1162) (1163) (1164) (1165) (1166) (1167) (1168) (1169) (1170) (1171) (1172) (1173) (1174) (1175) (1176) (1177) (1178) (1179) (1180) (1181) (1182) (1183) (1184) (1185) (1186) (1187) (1188) (1189) (1190) (1191) (1192) (1193) (1194) (1195) (1196) (1197) (1198) (1199) (1200) (1201) (1202) (1203) (1204) (1205) (1206) (1207) (1208) (1209) (1210) (1211) (1212) (1213) (1214) (1215) (1216) (1217) (1218) (1219) (1220) (1221) (1222) (1223) (1224) (1225) (1226) (1227) (1228) (1229) (1230) (1231) (1232) (1233) (1234) (1235) (1236) (1237) (1238) (1239) (1240) (1241) (1242) (1243) (1244) (1245) (1246) (1247) (1248) (1249) (1250) (1251) (1252) (1253) (1254) (1255) (1256) (1257) (1258) (1259) (1260) (1261) (1262) (1263) (1264) (1265) (1266) (1267) (1268) (1269) (1270) (1271) (1272) (1273) (1274) (1275) (1276) (1277) (1278) (1279) (1280) (1281) (1282) (1283) (1284) (1285) (1286) (1287) (1288) (1289) (1290) (1291) (1292) (1293) (1294) (1295) (1296) (1297) (1298) (1299) (1300) (1301) (1302) (1303) (1304) (1305) (1306) (1307) (1308) (1309) (1310) (1311) (1312) (1313) (1314) (1315) (1316) (1317) (1318) (1319) (1320) (1321) (1322) (1323) (1324) (1325) (1326) (1327) (1328) (1329) (1330) (1331) (1332) (1333) (1334) (1335) (1336) (1337) (1338) (1339) (1340) (1341) (1342) (1343) (1344) (1345) (1346) (1347) (1348) (1349) (1350) (1351) (1352) (1353) (1354) (1355) (1356) (1357) (1358) (1359) (1360) (1361) (1362) (1363) (1364) (1365) (1366) (1367) (1368) (1369) (1370) (1371) (1372) (1373) (1374) (1375) (1376) (1377) (1378) (1379) (1380) (1381) (1382) (1383) (1384) (1385) (1386) (1387) (1388) (1389) (1390) (1391) (1392) (1393) (1394) (1395) (1396) (1397) (1398) (1399) (1400) (1401) (1402) (1403) (1404) (1405) (1406) (1407) (1408) (1409) (1410) (1411) (1412) (1413) (1414) (1415) (1416) (1417) (1418) (1419) (1420) (1421) (1422) (1423) (1424) (1425) (1426) (1427) (1428) (1429) (1430) (1431) (1432) (1433) (1434) (1435) (1436) (1437) (1438) (1439) (1440) (1441) (1442) (1443) (1444) (1445) (1446) (1447) (1448) (1449) (1450) (1451) (1452) (1453) (1454) (1455) (1456) (1457) (1458) (1459) (1460) (1461) (1462) (1463) (1464) (1465) (1466) (1467) (1468) (1469) (1470) (1471) (1472) (1473) (1474) (1475) (1476) (1477) (1478) (1479) (1480) (1481) (1482) (1483) (1484) (1485) (1486) (1487) (1488) (1489) (1490) (1491) (1492) (1493) (1494) (1495) (1496) (1497) (1498) (1499) (1500) (1501) (1502) (1503) (1504) (1505) (1506) (1507) (1508) (1509) (1510) (1511) (1512) (1513) (1514) (1515) (1516) (1517) (1518) (1519) (1520) (1521) (1522) (1523) (1524) (1525) (1526) (1527) (1528) (1529) (1530) (1531) (1532) (1533) (1534) (1535) (1536) (1537) (1538) (1539) (1540) (1541) (1542) (1543) (1544) (1545) (1546) (1547) (1548) (1549) (1550) (1551) (1552) (1553) (1554) (1555) (1556) (1557) (1558) (1559) (1560) (1561) (1562) (1563) (1564) (1565) (1566) (1567) (1568) (1569) (1570) (1571) (1572) (1573) (1574) (1575) (1576) (1577) (1578) (1579) (1580) (1581) (1582) (1583) (1584) (1585) (1586) (1587) (1588) (1589) (1590) (1591) (1592) (1593) (1594) (1595) (1596) (1597) (1598) (1599) (1600) (1601) (1602) (1603) (1604) (1605) (1606) (1607) (1608) (1609) (1610) (1611) (1612) (1613) (1614) (1615) (1616) (1617) (1618) (1619) (1620) (1621) (1622) (1623) (1624) (1625) (1626) (1627) (1628) (1629) (1630) (1631) (1632) (1633) (1634) (1635) (1636) (1637) (1638) (1639) (1640) (1641) (1642) (1643) (1644) (1645) (1646) (1647) (1648) (1649) (1650) (1651) (1652) (1653) (1654) (1655) (1656) (1657) (1658) (1659) (1660) (1661) (1662) (1663) (1664) (1665) (1666) (1667) (1668) (1669) (1670) (1671) (1672) (1673) (1674) (1675) (1676) (1677) (1678) (1679) (1680) (1681) (1682) (1683) (1684) (1685) (1686) (1687) (1688) (1689) (1690) (1691) (1692) (1693) (1694) (1695) (1696) (1697) (1698) (1699) (1700) (1701) (1702) (1703) (1704) (1705) (1706) (1707) (1708) (1709) (1710) (1711) (1712) (1713) (1714) (1715) (1716) (1717) (1718) (1719) (1720) (1721) (1722) (1723) (1724) (1725) (1726) (1727) (1728) (1729) (1730) (1731) (1732) (1733) (1734) (1735) (1736) (1737) (1738) (1739) (1740) (1741) (1742) (1743) (1744) (1745) (1746) (1747) (1748) (1749) (1750) (1751) (1752) (1753) (1754) (1755) (1756) (1757) (1758) (1759) (1760) (1761) (1762) (1763) (1764) (1765) (1766) (1767) (1768) (1769) (1770) (1771) (1772) (1773) (1774) (1775) (1776) (1777) (1778) (1779) (1780) (1781) (1782) (1783) (1784) (1785) (1786) (1787) (1788) (1789) (1790) (1791) (1792) (1793) (1794) (1795) (1796) (1797) (1798) (1799) (1800) (1801) (1802) (1803) (1804) (1805) (1806) (1807) (1808) (1809) (1810) (1811) (1812) (1813) (1814) (1815) (1816) (1817) (1818) (1819) (1820) (1821) (1822) (1823) (1824) (1825) (1826) (1827) (1828) (1829) (1830) (1831) (1832) (1833) (1834) (1835) (1836) (1837) (1838) (1839) (1840) (1841) (1842) (1843) (1844) (1845) (1846) (1847) (1848) (1849) (1850) (1851) (1852) (1853) (1854) (1855) (1856) (1857) (1858) (1859) (1860) (1861) (1862) (1863) (1864) (1865) (1866) (1867) (1868) (1869) (1870) (1871) (1872) (1873) (1874) (1875) (1876) (1877) (1878) (1879) (1880) (1881) (1882) (1883) (1884) (1885) (1886) (1887) (1888) (1889) (1890) (1891) (1892) (1893) (1894) (1895) (1896) (1897) (1898) (1899) (1900) (1901) (1902) (1903) (1904) (1905) (1906) (1907) (1908) (1909) (1910) (1911) (1912) (1913) (1914) (1915) (1916) (1917) (1918) (1919) (1920) (1921) (1922) (1923) (1924) (1925) (1926) (1927) (1928) (1929) (1930) (1931) (1932) (1933) (1934) (1935) (1936) (1937) (1938) (1939) (1940) (1941) (1942) (1943) (1944) (1945) (1946) (1947) (1948) (1949) (1950) (1951) (1952) (1953) (1954) (1955) (1956) (1957) (1958) (1959) (1960) (1961) (1962) (1963) (1964) (1965) (1966) (1967) (1968) (1969) (1970) (1971) (1972) (1973) (1974) (1975) (1976) (1977) (1978) (1979) (1980) (1981) (1982) (1983) (1984) (1985) (1986) (1987) (1988) (1989) (1990) (1991) (1992) (1993) (1994) (1995) (1996) (1997) (1998) (1999) (2000) (2001) (2002) (2003) (2004) (2005) (2006) (2007) (2008) (2009) (2010) (2011) (2012) (2013) (2014) (2015) (2016) (2017) (2018) (2019) (2020) (2021) (2022) (2023) (2024) (2025) (2026) (2027) (2028) (2029) (2030) (2031) (2032) (2033) (2034) (2035) (2036) (2037) (2038) (2039) (2040) (2041) (2042) (2043) (2044) (2045) (2046) (2047) (2048) (2049) (2050) (2051) (2052) (2053) (2054) (2055) (2056) (2057) (2058) (2059) (2060) (2061) (2062) (2063) (2064) (2065) (2066) (2067) (2068) (2069) (2070) (2071) (2072) (2073) (2074) (2075) (2076) (2077) (2078) (2079) (2080) (2081) (2082) (2083) (2084) (2085) (2086) (2087) (2088) (2089) (2090) (2091) (2092) (2093) (2094) (2095) (2096) (2097) (2098) (2099) (2100) (2101) (2102) (2103) (2104) (2105) (2106) (2107) (2108) (2109) (2110) (2111) (2112) (2113) (2114) (2115) (2116) (2117) (2118) (2119) (2120) (2121) (2122) (2123) (2124) (2125) (2126) (2127) (2128) (2129) (2130) (2131) (2132) (2133) (2134) (2135) (2136) (2137) (2138) (2139) (2140) (2141) (2142) (2143) (2144) (2145) (2146) (2147) (2148) (2149) (2150) (2151) (2152) (2153) (2154) (2155) (2156) (2157) (2158) (2159) (2160) (2161) (2162) (2163) (2164) (2165) (2166) (2167) (2168) (2169) (2170) (2171) (2172) (2173) (2174) (2175) (2176) (2177) (2178) (2179) (2180) (2181) (2182) (2183) (2184) (2185) (2186) (2187) (2188) (2189) (2190) (2191) (2192) (2193) (2194) (2195) (2196) (2197) (2198) (2199) (2200) (2201) (2202) (2203) (2204) (2205) (2206) (2207) (2208) (2209) (2210) (2211) (2212) (2213) (2214) (2215) (2216) (2217) (2218) (2219) (2220) (2221) (2222) (2223) (2224) (2225) (2226) (2227) (2228) (2229) (2230) (2					



歯科レセプトの情報

- 医療機関毎・診療年月毎・入外毎にRE（レセプトヘッダ）が発生
- REに紐づく形で全15テーブル
- 研究に主として用いるテーブルは概ね8種類

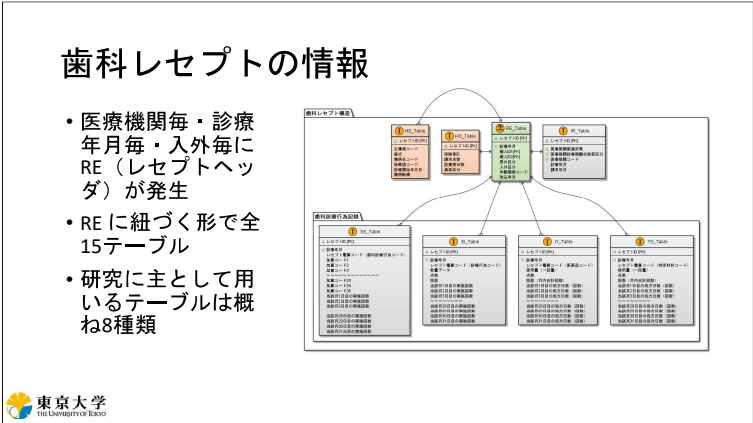
診療レセプトヘッダ

- 診療レセプト明細1
- 診療レセプト明細2
- 診療レセプト明細3
- 診療レセプト明細4

診療レセプト明細2

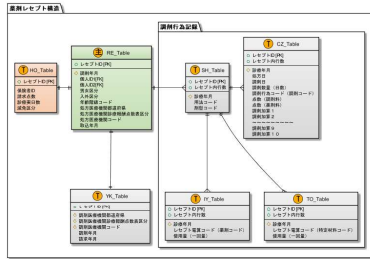
- 診療レセプト明細2-1
- 診療レセプト明細2-2
- 診療レセプト明細2-3
- 診療レセプト明細2-4

- # 歯科レセプトの情報
- 医療機関毎・診療年月毎・入外毎にRE（レセプトヘッダ）が発生
 - REに紐づく形で全15テーブル
 - 研究に主として用いるテーブルは概ね8種類
-
- 診療レセプトヘッダ
- 診療レセプト明細1
 - 診療レセプト明細2
 - 診療レセプト明細3
 - 診療レセプト明細4
- 診療レセプト明細2
- 診療レセプト明細2-1
 - 診療レセプト明細2-2
 - 診療レセプト明細2-3
 - 診療レセプト明細2-4



調剤レセプトの情報

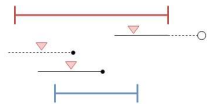
- 調剤機関毎・調剤年月毎にREが発生
- REに紐づく形で全14テーブル
- 研究に主として用いるテーブルは概ね7種類
- 薬剤の処方状況を確認するために、レセ中の出現行数を基にSHに接続が必要



オンサイトセンターでの研究の実際

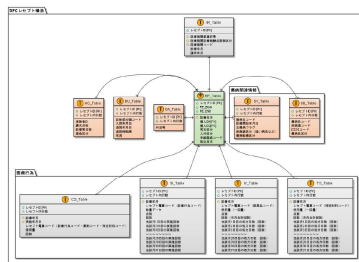
症例抽出計画の策定

- Inclusion Period の決定
 - 許諾された全期間で症例を検索する。
 - 容量削減や抽出プロセス短縮のために、期間を絞ることはある。
- キーレセプト条件は単一レセプト内で完結する条件にする。
 - 他のレセプトを参照して決定する条件は好ましくない。
 - 例えば、“脳梗塞病名のついた症例のレセプト”や“rPA実施症例のレセプト”は検索できる。
 - “脳梗塞発症後”、“リハビリテーション”を行ったレセプトの検索は危険
- 研究に利用する情報を決める。



DPCレセプトの情報

- 医療機関毎・診療年月毎に**複数**のREが発生
- REとその出現行数に紐づく形で全26テーブル
- 研究に主として用いるテーブルは概ね11種類



オンサイトセンターでの研究の実際

サンプルクエリ

```
以下ではtbl_ssz2 Extract にSSZ2 船を重複なく抽出するためのクエリを作成する。
以下参照
SELECT distinct_ssz2
INSERT INTO tbl_ssz2 Extract
SELECT DISTINCT SSZ2.NO, PRAC, TM FROM onsite.tbl_ssz2 WHERE PRAC, TM = '202110' AND (random() < 0.05)
UNION
SELECT DISTINCT SSZ2.NO, PRAC, TM FROM onsite.tbl_ssz2 WHERE PRAC, TM = '202110' AND (random() < 0.05)
LIMIT 10000
SELECT SSZ2.NO FROM onsite.tbl_ssz2 Extract LIMIT 10000 OFFSET 10000
```

演習（NDBの構造）の演習4に取り組んでください。

NDBデータ構造に触れる

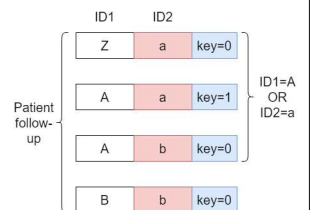
- 学習を円滑に進めるため、一時的なWebサイトを公開。
- 演習1にアクセスしてください。

準備中

オンサイトセンターでの症例の実際

症例の追跡

- 症例追跡期間の決定
 - 許諾された全期間でキーレセプト前後情報を追跡する。
- 症例の追跡
 - キーレセプトを基にID1, ID2を再帰的に追跡する。
 - 'ID1 = OR ID2 = 'の条件一回で追跡を打ち切るのはダメ
- 追跡したID1, ID2を基にID0を作成する。
 - 別人を誤って追跡した場合は後でその症例を除くが稀である。
 - ランダム抽出する場合は、ID0単位で抽出が望ましい。
- 演習（NDBの構造）の演習10に取り組んでください。



オンサイトセンターでの研究の実際

- 症例抽出計画の策定
 - Inclusion Period の決定
 - 症例抽出条件の（キーレセプトの抽出）決定
 - 抽出する情報の決定
- 症例の追跡
 - 症例追跡期間の決定
- レセプト情報の整理（コホート化）
- 解析

オンサイトセンターでの症例の実際

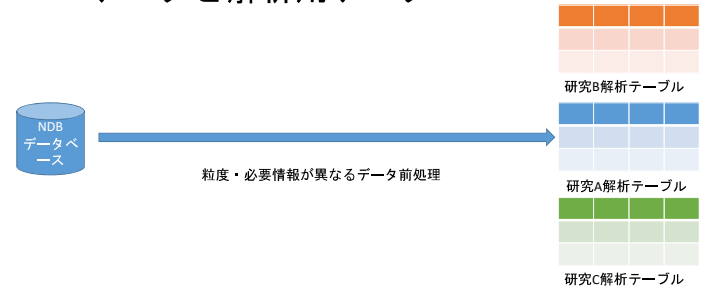
レセプト情報の整理（コホート化）

- 対象となるIDを基にレセプトを検索
- レセプトから必要となる情報を抽出する。
 - 実施医療行為や傷病名に加え特定健診情報等
 - 情報抽出過程はレセプト単位で独立させる。
 - MEDISの標準マスタなどを活用
- 時系列に情報を並べる（解析データベースの構築・コホート化）。

解析用データセットを設計する

- NDBに含まれるデータとその構造を学習した。
- NDBデータは構造として使いにくい
 - ・情報がテーブル間にまたがって存在
 - ・テーブルの中は月単位情報と日単位情報がバラバラに
 - ・レセプト単位・処置単位でとれる情報が一つのテーブルに
- 中級編で触れたような、解析しやすい形のデータセットを作成したい。
- 設計の考え方を学ぶ

NDBデータと解析用データ



解析用データセットを設計する

NDBデータベースに格納されている情報
DBから得られるタイムスタンプ（時間軸の粒度）

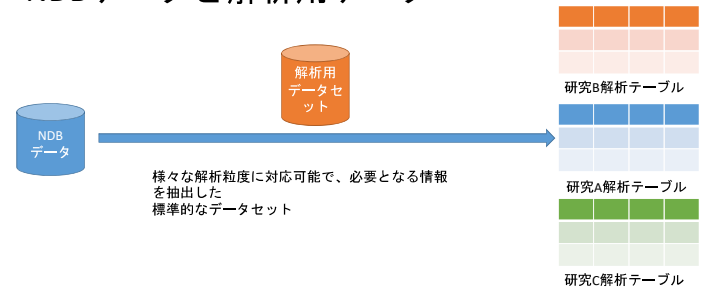


DBから得られる集計単位



DBから得られる情報

NDBデータと解析用データ



解析用データセットを設計する。

- レセプトデータはテーブル構造が複雑で解析しにくい
- 解析しやすいデータとは？
- 収集用（格納用）のデータ構造と解析用のデータ構造は異なる。
- R, Stata, SPSS, SAS, Python いずれも、基本的に1つの表（テーブル）に情報を集約する。
 - ・行：解析粒度
 - ・列：情報

解析粒度（テーブル1行）

患者ID	年齢	性別	治療薬1		転帰	最終観察

データ次元

解析用データセットの設計

- NDBデータに含まれる情報（エンティティ）をリストアップする。
- 各エンティティに紐づくの属性を整理する。
- 各エンティティの時間粒度を整理する。
- 各エンティティ間の接続・依存関係（1:N, 1:1, N:M）を整理する。
- 時間粒度と紐づく属性が似ている情報を1つのテーブルにまとめる。
 - ・この際、データベース正規化をどこまで行うかは個別事例による。

データハンドリング

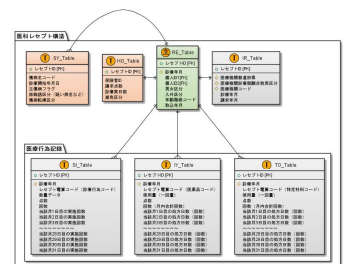
- 必要な解析粒度は解析手法によって異なる。
 - ・患者単位
 - ・入院単位
 - ・地域単位
 - ・病院単位
 - ・期間単位
 - ・患者日単位
- 必要なデータ次元も研究によって異なる。
 - ・特定治療薬・特定診療行為の有無
 - ・特定診断名

解析粒度（テーブル1行）

患者ID	年齢	性別	治療薬1		転帰	最終観察

医療レセプトの情報（エンティティ）

- 医療機関毎・診療年月毎・入外毎にRE（レセプトヘッダ）が発生
- RE に紐づく形で全18テーブル
- 研究に主として用いるテーブルは概ね7種類

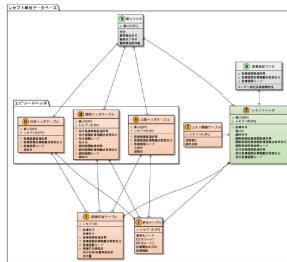


医科レセプトの情報（エンティティ）

- 患者
- 最終転帰
- レセプト発生
- 患者の外来受診
- 患者の入院
- 診断
- 診療行為記録（診療行為・医薬品・特定材料）

解析用データベース構造の一例

- 当教室で解析用DBとして用いているData Model



医科レセプトエンティティに紐づく属性

- 患者
 - 患者ID・年齢・性別・観察開始年月日・観察終了年月日
- 最終転帰
 - 患者ID・死亡転帰
- レセプト発生
 - 患者ID・レセプトID・診療年月・月額保険費用・保険者区分・レセプト種別
- 患者の外来受診
 - レセプトID・外来受診日・医療機関
- 患者の入院
 - レセプトID・入院日・退院日・医療機関
- 診断
 - レセプトID・診療年月・傷病名区分コード・修飾語コード・診療開始年月日
- 診療行為記録（診療行為・医薬品・特定材料）
 - レセプトID・レセプト電算コード・実施日・使用量・回数・点数

解析用データベースから研究用テーブルを作る

レセプトを跨いだ症例情報を取得する（研究用テーブル）。

- Date Zero の設定
- Washout Period の設定
- Follow-up と打ち切りの設定
- Exposure の設定
- Outcome の設定
- Covariate の設定
- 中級編で学んだこと

医科レセプトエンティティ時間粒度

- 患者
 - 患者ID・年齢・性別・観察開始年月日・観察終了年月日
- 最終転帰
 - 患者ID・死亡転帰・死亡観察日
- レセプト発生
 - 患者ID・レセプトID・診療年月・月額保険費用・保険者区分・レセプト種別
- 患者の外来受診
 - レセプトID・外来受診日・医療機関
- 患者の入院
 - レセプトID・入院日・退院日・医療機関
- 診断
 - レセプトID・診療年月・傷病名区分コード・修飾語コード・診療開始年月日
- 診療行為記録（診療行為・医薬品・特定材料）
 - レセプトID・レセプト電算コード・実施日・使用量・回数・点数

解析データベースを構築する

- 学習を円滑に進めるため、一時的なWebサイトを公開。
- 参考のページにアクセスしてください。

準備中

用意するテーブル

過去に実績のある研究用データセットの仕様

- 患者IDテーブル
 - 患者個人に対し1行
- レセプトヘッダテーブル
 - レセプト（診療年月）ごとに1行（DPCレセプトについては複数行）
- 医療行為テーブル
 - 診療行為・医薬品・材料の日ごとの実施状況
- 診断テーブル
 - 月×診療開始年月日毎（DPCレセプトについては入院日毎）の診断情報
- アウトカムテーブル
 - 死亡転帰

まとめ

大規模なレセプトデータベースを解析する場合に当たって、

- レセプトデータベースの構造を学んだ。
- 解析用データセットの設計・構築の流れを学習した。

- 演習用ページを利用して、理解をさらに深めてください。