

공개형 데이터 설명서

대장암 다기관 암 임상 라이브러리 합성데이터

- 정밀의료 구현을 위한 다기관적 암환자 합성 의료 데이터의 생성 및 응용연구, 보건산업진흥원 -

2023. 12. 20.



□ 개요 :

- 인공지능 기술의 비약적인 발전으로 인해 사회적·국가적 투자가 전략적으로 이루어지고 있으며, 국내에서도 임상데이터 기반 인공지능 기술의 선도적인 거점 확보를 위해 많은 연구 투자가 진행되고 있으나, 원천데이터 사용의 제약성으로 인해 개방 및 활용성이 매우 미비함
- 이들 문제해결을 위해 합성데이터 생성 알고리즘을 적용하여 생성한 가상 데이터인 대장암 임상라이브러리 합성데이터를 개방함으로써 인공지능 기술개발 활성화 및 확인을 기대함

□ 데이터 구성

- 대장암 합성 데이터는 암 등록, 건강정보, 신체계측정보, 면역병리, 분자병리, 수술여부, 항암제치료 등으로 구성

- 삼성서울병원 (총 123개 변수)

번호	항목명	테이블명	컬럼수	데이터수
1	대장암_환자_기본정보	clrc_pt_bsnf	19	52767
2	대장암_검사_진단	clrc_ex_diag	9	520712
3	대장암_병리_생체검사	clrc_pth_bpsy	8	77981
4	대장암_병리_면역	clrc_pth_mnty	14	28449
5	대장암_병리_분자	clrc_pth_mlcrr	16	39893
6	대장암_병리_외과	clrc_pth_srgc	26	13593
7	대장암_수술_정보	clrc_oprt_nfrm	9	35454
8	대장암_치료_항암제	clrc_trtm_casb	12	113919
9	대장암_치료_방사선	clrc_trtm_rd	6	9797
10	대장암_사망_정보	clrc_dead_nfrm	4	11219

○ 대장암 임상 라이브러리 합성데이터 코드북 (GAN)

- 연세대학교병원 (총 131개 변수)

번호	항목명	테이블명	컬럼수	데이터수
1	대장암_환자_기본정보	clrc_pt_bsnf	19	54332
2	대장암_검사_진단	clrc_ex_diag	9	525872
3	대장암_병리_생체검사	clrc_ptb_bpsy	10	24899
4	대장암_병리_면역	clrc_ptb_mnty	14	3556
5	대장암_병리_분자	clrc_ptb_mlc	16	13983
6	대장암_병리_외과	clrc_ptb_srgc	26	46047
7	대장암_수술_정보	clrc_oprt_nfrm	9	44993
8	대장암_치료_항암제	clrc_trtm_casb	12	213210
9	대장암_치료_방사선	clrc_trtm_rd	8	1094
10	대장암_사망_정보	clrc_dead_nfrm	4	5416
11	대장암_진단_재발정보	clrc_dg_rcnf	4	3071

- 국립암센터 (총 117개 변수)

번호	항목명	테이블명	컬럼수	데이터수
1	대장암_환자_기본정보	clrc_pt_bsnf	19	54312
2	대장암_검사_진단	clrc_ex_diag	9	460482
3	대장암_병리_생체검사	clrc_ptb_bpsy	8	588090
4	대장암_병리_면역	clrc_ptb_mnty	14	2890
5	대장암_병리_분자	clrc_ptb_mlc	16	12566
6	대장암_병리_외과	clrc_ptb_srgc	20	27612
7	대장암_수술_정보	clrc_oprt_nfrm	9	23549
8	대장암_치료_항암제	clrc_trtm_casb	12	64816
9	대장암_치료_방사선	clrc_trtm_rd	6	2572
10	대장암_사망_정보	clrc_dead_nfrm	4	1697

- 삼성서울병원

중분류 (Table)	변수명 (Column)	변수 설명	변수값(코드) 설명 (99=모름/기타/미측정)
대장암_환자_기본정보	CENTER_CD	센터코드	센터코드 (2자리)
대장암_환자_기본정보	PT_SBST_NO	환자대체번호	개인고유번호 예) X000000003
대장암_환자_기본정보	BSPT_SEX_CD	기본환자성별코드	환자의 성별 M : 남자 F : 여자 E : 기타
대장암_환자_기본정보	BSPT_BRYM	기본환자생년월	생년월 YYYYMM
대장암_환자_기본정보	BSPT_FRST_DIAG_YMD	기본환자최초진단일자	대장암초진일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_FRST_DIAG_CD	기본환자최초진단코드	암진단코드 예) C18
대장암_환자_기본정보	BSPT_FRST_DIAG_NM	기본환자최초진단명	암진단명 예) Malignant neoplasm of colon
대장암_환자_기본정보	BSPT_IDGN_AGE	기본환자진단시연령	대장암 최초 진단시 나이(만) 예) 45 ※주진단으로 C18-20(하위코드포함)를 받은 환자
대장암_환자_기본정보	BSPT_FRST_OPRT_YMD	기본환자최초수술일자	최초 대장암 진단 이후 첫수술일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_FRST_ANCN_TRT_M_STRT_YMD	기본환자최초항암치료시작일자	최초항암치료시작일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_FRST_RDT_STRT_YMD	기본환자최초방사선치료시작일자	최초 대장암 진단 이후 첫 방사선치료시작일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_DEAD_YMD	기본환자사망일자	사망일자 YYYY-MM-DD
대장암_환자_기본정보	CENTER_LAST_VST_YMD	센터 최종방문일자	센터 최종 방문일 YYYY-MM-DD
대장암_환자_기본정보	OVRL_SRVL_DTRN_DCNT	전체생존기간일수	환자의 진단일부터 사망일 또는 최종내원일까지의 기간(일)
대장암_환자_기본정보	BSPT_STAG_VL	기본환자병기값	기본환자병기값 예) 0
대장암_환자_기본정보	BSPT_T_STAG_VL	기본환자T병기값	기본환자T병기값 예) 0

대장암_환자_기본정보	BSPT_N_STAG_VL	기본환자N병기값	기본환자N병기값 예) 0
대장암_환자_기본정보	BSPT_M_STAG_VL	기본환자M병기값	기본환자M병기값 예) 0
대장암_환자_기본정보	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_검사_진단	CENTER_CD	센터코드	센터코드 (2자리)
대장암_검사_진단	PT_SBST_NO	환자대체번호	개인고유번호 예) X000000003
대장암_검사_진단	CEXM_YMD	진단검사일자	진단검사일자 (수집기준일자) YYYY-MM-DD
대장암_검사_진단	CEXM_SEQ	진단검사순번	진단검사순번 (정수) 예) 1
대장암_검사_진단	CEXM_NM	진단검사명	진단검사명 예) CEA
대장암_검사_진단	CEXM_RSLT_CONT	진단검사결과내용	진단검사 결과 내용 예) 33.2
대장암_검사_진단	CEXM_RSLT_UNIT_CON NT	진단검사결과단위내용	진단검사 결과 내용의 단위 예) ng/ml
대장암_검사_진단	CEXM_RSLT_REF_CON T	진단검사결과참고내용	진단검사 종류 참고치 예) 0~5.0
대장암_검사_진단	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_생체검사	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_생체검사	PT_SBST_NO	환자대체번호	개인고유번호 예) X000000003
대장암_병리_생체검사	BPTH_ACPT_YMD	생체검사병리접수일자	생검 접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_생체검사	BPTH_SEQ	생체검사병리순번	생검 순번 (정수) 예) 1
대장암_병리_생체검사	BPTH_READ_YMD	생체검사병리판독일자	생검 판독일자 YYYY-MM-DD
대장암_병리_생체검사	BPTH_SITE_CONT	생체검사병리부위내용	시행된 생체검사 부위 예) colon
대장암_병리_생체검사	BPTH_BPSY_RSLT_CON T	생체검사병리생체검사결과내용	생체검사 결과 내용 예) carcinoma
대장암_병리_생체검사	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_면역	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_면역	PT_SBST_NO	환자대체번호	개인고유번호 예) X000000003
대장암_병리_면역	IMPT_ACPT_YMD	면역병리접수일자	면역병리검사접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_면역	IMPT_SEQ	면역병리순번	면역병리 정수 순번 예) 1

대장암_병리_면역	IMPT_READ_YMD	면역병리판독일자	면역병리판독일자 YYYY-MM-DD
대장암_병리_면역	IMPT_HM1E_RSLT_CD	면역병리HMLH1검사결과 코드	hMLH1 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HM1E_RSLT_NM	면역병리HMLH1검사결과 명	hMLH1 결과 명 예) loss
대장암_병리_면역	IMPT_HS2E_RSLT_CD	면역병리HMSH2검사결과 코드	hMLH2 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HS2E_RSLT_NM	면역병리HMSH2검사결과 명	hMLH2 결과 명 예) loss
대장암_병리_면역	IMPT_HS6E_RSLT_CD	면역병리HMSH6검사결과 코드	hMLH6 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HS6E_RSLT_NM	면역병리HMSH6검사결과 명	HMSH6 결과 명 예) loss
대장암_병리_면역	IMPT_HP2E_RSLT_CD	면역병리HPMS2검사결과 코드	hPMS2 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HP2E_RSLT_NM	면역병리HPMS2검사결과 명	HPMS2 결과 명 예) loss
대장암_병리_면역	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_분자	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_분자	PT_SBST_NO	환자대체번호	개인고유번호 예) X000000003
대장암_병리_분자	MLPT_ACPT_YMD	분자병리접수일자	분자병리검사접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_분자	MLPT_SEQ	분자병리순번	분자병리순번 예) 1
대장암_병리_분자	MLPT_READ_YMD	분자병리판독일자	면역병리판독일자 YYYY-MM-DD
대장암_병리_분자	MLPT_MSIE_RSLT_CD	분자병리MSI검사결과코드	MSI검사결과코드 1 : stable 2 : low 3 : high 9 : Other
대장암_병리_분자	MLPT_MSIE_RSLT_NM	분자병리MSI검사결과명	MSI검사결과명 예) stable

대장암_병리_분자	MLPT_KE2E_RSLT_CD	분자병리KRASMUTATION_EXON2검사결과코드	K-ras mutation (exon2) 결과 코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_KE2E_RSLT_NM	분자병리KRASMUTATION_EXON2검사결과명	K-ras mutation (exon2) 결과명 예) not detected
대장암_병리_분자	MLPT_KRES_RSLT_CD	분자병리KRASMUTATION 검사결과코드	K-ras mutation 결과 코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_KRES_RSLT_NM	분자병리KRASMUTATION 검사결과명	K-ras mutation결과명 예) not detected
대장암_병리_분자	MLPT_NREX_RSLT_CD	분자병리NRASMUTATION 검사결과코드	N-ras mutation 결과코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_NREX_RSLT_NM	분자병리NRASMUTATION 검사결과명	N-ras mutation 결과명 예) detected
대장암_병리_분자	MLPT_BRME_RSLT_CD	분자병리BRAF_MUTATIO N검사결과코드	BRAF mutation 결과코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_BRME_RSLT_NM	분자병리BRAF_MUTATIO N검사결과명	BRAF mutation 결과명 예) detected
대장암_병리_분자	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_외과	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_외과	PT_SBST_NO	환자대체번호	개인고유번호 예) X000000003
대장암_병리_외과	SGPT_ACPT_YMD	외과병리접수일자	외과병리접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_외과	SGPT_SEQ	외과병리순번	외과병리순번 (정수) 예) 1
대장암_병리_외과	SGPT_READ_YMD	외과병리판독일자	외과병리판독일자 YYYY-MM-DD
대장암_병리_외과	SGPT_PATL_STAG_VL	외과병리병리학적병기값	외과병리병리학적병기값 예)pT3N2bM1a
대장암_병리_외과	SGPT_PATL_T_STAG_VL	외과병리병리학적T병기값	외과병리병리학적T병기값 예)3a
대장암_병리_외과	SGPT_PATL_N_STAG_VL	외과병리병리학적N병기값	외과병리병리학적N병기값 예)2b

대장암_병리_외과	SGPT_PATL_M_STAG_VL	외과병리병리학적M병기값	외과병리병리학적M병기값 예)1a
대장암_병리_외과	SGPT_SRMV_LN_CNT	외과병리절제림프절수	절제한 전체 림프절 개수 예)44
대장암_병리_외과	SGPT_MTST_LN_CNT	외과병리전이림프절수	절제한 림프절 중 전이된 림프절 개수 예)1
대장암_병리_외과	SGPT_SRMG_PCTS_STA T_CD	외과병리수술절제면근위 암조직크기상태코드	수술절제면에 암조직 잔존유무 상태 코드 (proximal) 1: involed 2: uninvoled 3: indeterminate 9: Other
대장암_병리_외과	SGPT_SRMG_PCTS_STA T_NM	외과병리수술절제면근위 암조직크기상태명	수술절제면에 암조직 잔존유무 상태 명 (proximal) 예) involed
대장암_병리_외과	SGPT_SRMG_DCTS_ST AT_CD	외과병리수술절제면원위 암조직크기상태코드	수술절제면에 암조직 잔존유무 상태 코드 (distal) 1: involed 2: uninvoled 3: indeterminate 9: Other
대장암_병리_외과	SGPT_SRMG_DCTS_ST AT_NM	외과병리수술절제면원위 암조직크기상태명	수술절제면에 암조직 잔존유무 상태 명 (distal) 예) involed
대장암_병리_외과	SGPT_SRMG_RCTS_ST AT_CD	외과병리수술절제면방사 형암조직크기상태코드	수술절제면에 암조직 잔존유무 상태 코드 (radial(rectum only)) 1: involed 2: uninvoled 3: indeterminate 9: Other
대장암_병리_외과	SGPT_SRMG_RCTS_ST AT_NM	외과병리수술절제면방사 형암조직크기상태명	수술절제면에 암조직 잔존유무 상태 명 (radial(rectum only)) 예) involed
대장암_병리_외과	SGPT_NERV_PREX_CD	외과병리신경주위침윤코드	신경주위침윤코드 1: present 2: absent 3: not identified 4: no record 9: Other
대장암_병리_외과	SGPT_NERV_PREX_NM	외과병리신경주위침윤명	신경주위침윤명 예) present
대장암_병리_외과	SGPT_VNIN_CD	외과병리정맥침윤코드	정맥침윤상태코드 1: present 2: absent 3: not identified 4: no record 9: Other
대장암_병리_외과	SGPT_VNIN_NM	외과병리정맥침윤명	정맥침윤상태명 예) present

대장암_병리_외과	SGPT_ANIN_CD	외과병리혈관조영침윤코드	혈관조영침윤코드 1: present 2: absent 3: not identified 4: no record 9: Other
대장암_병리_외과	SGPT_ANIN_NM	외과병리혈관조영침윤명	혈관조영침윤명 예) present
대장암_병리_외과	SGPT_TUMR_BUDD_CD	외과병리종양발아코드	Tumor budding 코드 1: present 2: absent 3: not identified 4: no record 9: Other
대장암_병리_외과	SGPT_TUMR_BUDD_NM	외과병리종양발아명	Tumor budding 명 예) present
대장암_병리_외과	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_수술_정보	CENTER_CD	센터코드	센터코드 (2자리)
대장암_수술_정보	PT_SBST_NO	환자대체번호	개인고유번호 예) X000000003
대장암_수술_정보	OPRT_YMD	수술일자	수술일자 (수집기준일자) YYYY-MM-DD
대장암_수술_정보	OPRT_SEQ	수술순번	수술순번 (정수) 예) 1
대장암_수술_정보	OPRT_CLCN_OPRT_KIN_D_CD	수술대장암수술종류코드	대장암수술종류코드 01 : 결장절제술 02 : 반흔허니아근본수술 03 : 소장절제술 04 : 장 및 장간막 손상수술 05 : 장간막종양적출술 06 : 장관유착박리술 07 : 장루조성술 08 : 장루폐쇄술 09 : 장문합술 10 : 장폐색증수술 11 : 직장 및 에스장절제술 12 : 직장종양 절제술 13 : 직장항문 주위농양수술 14 : 진단적개복술 15 : 충수절제술 16 : 후복막종양적출술 99 : 기타
대장암_수술_정보	OPRT_CLCN_OPRT_KIN_D_NM	수술대장암수술종류명	대장암수술종류코드명 예) 결장절제술

대장암_수술_정보	OPRT_CURA_RSCT_CD	수술근치적절제술코드	Curative resection 코드 1: 해당없음 2: curative 3: palliative 9: Other
대장암_수술_정보	OPRT_CURA_RSCT_NM	수술근치적절제술명	Curative resection명 예) curative
대장암_수술_정보	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_치료_항암제	CENTER_CD	센터코드	센터코드 (2자리)
대장암_치료_항암제	PT_SBST_NO	환자대체번호	개인고유번호 예) X000000003
대장암_치료_항암제	CSTR_STRT_YMD	항암제치료시작일자	일자 YYYY-MM-DD
대장암_치료_항암제	CSTR_SEQ	항암제치료순번	항암제치료방순번 (정수) 예) 1
대장암_치료_항암제	CSTR_REGN_CD	항암제치료요법코드	Regimen 코드 01: none 02 :5-FU ± leucovorin 03 : capecitabine 04 : S-1 05 : UFT ± leucovorin 06 : avastin + fluoropyrimidine(FL/X/UFT/S-1) 07 : FOLFOX 08 : FOLFOX + AVASTIN 09 : FOLFOX + ERBITUX 10 : XELOX 11 : XELOX + AVASTIN 12 : XELOX + ERBITUX 13 : FOLFIRI 14 : FOLFIRI + AVASTIN 15 : FOLFIRI + ERBITUX 99 : Other
대장암_치료_항암제	CSTR_REGN_NM	항암제치료요법명	Regimen 명 예) S-1
대장암_치료_항암제	CSTR_NT	항암제치료횟수	항암제치료횟수(Line) 예) 10
대장암_치료_항암제	CSTR_END_YMD	항암제치료종료일자	일자 YYYY-MM-DD

대장암_치료_항암제	CSTR_PRPS_CD	항암제치료목적코드	항암치료목적코드 01: Neo-adjuvant 02: Adjuvant 03: Palliative 04: Concurrent 05: Induction 06: Maintenance 07: Salvage 08: Consolidation 99: Other
대장암_치료_항암제	CSTR_PRPS_NM	항암제치료목적명	항암치료목적명 예) Adjuvant
대장암_치료_항암제	CSTR_CYCL_VL	항암제치료주기값	항암치료주기값 예) 5
대장암_치료_항암제	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_치료_방사선	CENTER_CD	센터코드	센터코드 (2자리)
대장암_치료_방사선	PT_SBST_NO	환자대체번호	개인고유번호 예) X000000003
대장암_치료_방사선	RDT_STRT_YMD	방사선치료시작일자	방사선치료시작일자 (수집기준일자) YYYY-MM-DD
대장암_치료_방사선	RDT_RT_NO	방사선치료RT번호	방사선치료RT번호 예) 1
대장암_치료_방사선	RDT_END_YMD	방사선치료종료일자	일자 YYYY-MM-DD
대장암_치료_방사선	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_사망_정보	CENTER_CD	센터코드	센터코드 (2자리)
대장암_사망_정보	PT_SBST_NO	환자대체번호	개인고유번호 예) X000000003
대장암_사망_정보	DEAD_YMD	사망일자	사망일자 (수집기준일자) YYYY-MM-DD
대장암_사망_정보	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()

- 연세대학교병원

중분류 (Table)	변수명 (Column)	변수 설명	변수값(코드) 설명 (99=모름/기타/미측정)
대장암_환자_기본정보	CENTER_CD	센터코드	센터코드 (2자리)
대장암_환자_기본정보	PT_SBST_NO	환자대체번호	개인고유번호 예) RN000000002

대장암_환자_기본정보	BSPT_SEX_CD	기본환자성별코드	환자의 성별 M : 남자 F : 여자 E : 기타
대장암_환자_기본정보	BSPT_BRYM	기본환자생년월	생년월 YYYYMM
대장암_환자_기본정보	BSPT_FRST_DIAG_YMD	기본환자최초진단일자	대장암초진일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_FRST_DIAG_CD	기본환자최초진단코드	암진단코드 예) C18
대장암_환자_기본정보	BSPT_FRST_DIAG_NM	기본환자최초진단명	암진단명 예) colon
대장암_환자_기본정보	BSPT_IDGN_AGE	기본환자진단시연령	대장암 최초 진단시 나이(만) 예) 45 ※주진단으로 C18-20(하위코드포함)를 받은 환자
대장암_환자_기본정보	BSPT_FRST_OPRT_YMD	기본환자최초수술일자	최초 대장암 진단 이후 첫수술일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_FRST_ANCN_TRT_M_STRT_YMD	기본환자최초항암치료시작일자	최초항암치료시작일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_FRST_RDT_STRT_YMD	기본환자최초방사선치료시작일자	최초 대장암 진단 이후 첫 방사선치료시작일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_DEAD_YMD	기본환자사망일자	사망일자 YYYY-MM-DD
대장암_환자_기본정보	CENTER_LAST_VST_YMD	센터최종방문일자	센터 최종 방문일 YYYY-MM-DD
대장암_환자_기본정보	OVRL_SRLV_DTRN_DCNT	전체생존기간일수	환자의 진단일부터 사망일 또는 최종내원일까지의 기간(일)
대장암_환자_기본정보	BSPT_STAG_VL	기본환자병기값	기본환자병기값 예) 0
대장암_환자_기본정보	BSPT_T_STAG_VL	기본환자T병기값	기본환자T병기값 예) 0
대장암_환자_기본정보	BSPT_N_STAG_VL	기본환자N병기값	기본환자N병기값 예) 0
대장암_환자_기본정보	BSPT_M_STAG_VL	기본환자M병기값	기본환자M병기값 예) 0
대장암_환자_기본정보	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_검사_진단	CENTER_CD	센터코드	센터코드 (2자리)
대장암_검사_진단	PT_SBST_NO	환자대체번호	개인고유번호 예) RN000000002
대장암_검사_진단	CEXM_YMD	진단검사일자	진단검사일자 (수집기준일자) YYYY-MM-DD

대장암_검사_진단	CEXM_SEQ	진단검사순번	진단검사순번 (정수) 예) 1
대장암_검사_진단	CEXM_NM	진단검사명	진단검사명 예) CEA
대장암_검사_진단	CEXM_RSLT_CONT	진단검사결과내용	진단검사 결과 내용 예) 33.2
대장암_검사_진단	CEXM_RSLT_UNIT_CO NT	진단검사결과단위내용	진단검사 결과 내용의 단위 예) ng/ml
대장암_검사_진단	CEXM_RSLT_REF_CON T	진단검사결과참고내용	진단검사 종류 참고치 예) 0~5.0
대장암_검사_진단	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_생체검 사	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_생체검 사	PT_SBST_NO	환자대체번호	개인고유번호 예) RN000000002
대장암_병리_생체검 사	BPTH_ACPT_YMD	생체검사병리접수일자	생검 접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_생체검 사	BPTH_SEQ	생체검사병리순번	생검 순번 (정수) 예) 1
대장암_병리_생체검 사	BPTH_READ_YMD	생체검사병리판독일자	생검 판독일자 YYYY-MM-DD
대장암_병리_생체검 사	BPTH_SITE_CONT	생체검사병리부위내용	시행된 생체검사 부위 예) colon
대장암_병리_생체검 사	BPTH_BPSY_RSLT_CON T	생체검사병리생체검사결 과내용	생체검사 결과 내용 예) carcinoma
대장암_병리_생체검 사	BPTH_CELL_DIFF_CD	생체검사병리세포분화코 드	Cellular differentiation코드 01: Adenocarcinoma, WD 02: Adenocarcinoma, MD 03: Adenocarcinoma, PD 04: Mucinous ca 05: Carcinoid 99: Other
대장암_병리_생체검 사	BPTH_CELL_DIFF_NM	생체검사병리세포분화명	Cellular differentiation명 예) Adenocarcinoma, WD
대장암_병리_생체검 사	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_면역	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_면역	PT_SBST_NO	환자대체번호	개인고유번호 예) RN000000002
대장암_병리_면역	IMPT_ACPT_YMD	면역병리접수일자	면역병리검사접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_면역	IMPT_SEQ	면역병리순번	면역병리 정수 순번 예) 1

대장암_병리_면역	IMPT_READ_YMD	면역병리판독일자	면역병리판독일자 YYYY-MM-DD
대장암_병리_면역	IMPT_HM1E_RSLT_CD	면역병리HMLH1검사결과 코드	hMLH1 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HM1E_RSLT_NM	면역병리HMLH1검사결과 명	hMLH1 결과 명 예) loss
대장암_병리_면역	IMPT_HS2E_RSLT_CD	면역병리HMSH2검사결과 코드	hMLH2 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HS2E_RSLT_NM	면역병리HMSH2검사결과 명	hMLH2 결과 명 예) loss
대장암_병리_면역	IMPT_HS6E_RSLT_CD	면역병리HMSH6검사결과 코드	hMLH6 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HS6E_RSLT_NM	면역병리HMSH6검사결과 명	HMSH6 결과 명 예) loss
대장암_병리_면역	IMPT_HP2E_RSLT_CD	면역병리HPMS2검사결과 코드	hPMS2 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HP2E_RSLT_NM	면역병리HPMS2검사결과 명	HPMS2 결과 명 예) loss
대장암_병리_면역	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_분자	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_분자	PT_SBST_NO	환자대체번호	개인고유번호 예) RN000000002
대장암_병리_분자	MLPT_ACPT_YMD	분자병리접수일자	분자병리검사접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_분자	MLPT_SEQ	분자병리순번	분자병리순번 예) 1
대장암_병리_분자	MLPT_READ_YMD	분자병리판독일자	면역병리판독일자 YYYY-MM-DD
대장암_병리_분자	MLPT_MSIE_RSLT_CD	분자병리MSI검사결과코드	MSI검사결과코드 1 : stable 2 : low 3 : high 9 : Other
대장암_병리_분자	MLPT_MSIE_RSLT_NM	분자병리MSI검사결과명	MSI검사결과명 예) stable

대장암_병리_분자	MLPT_KE2E_RSLT_CD	분자병리KRASMUTATION_EXON2검사결과코드	K-ras mutation (exon2) 결과 코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_KE2E_RSLT_NM	분자병리KRASMUTATION_EXON2검사결과명	K-ras mutation (exon2) 결과명 예) not detected
대장암_병리_분자	MLPT_KRES_RSLT_CD	분자병리KRASMUTATION 검사결과코드	K-ras mutation 결과 코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_KRES_RSLT_NM	분자병리KRASMUTATION 검사결과명	K-ras mutation결과명 예) not detected
대장암_병리_분자	MLPT_NREX_RSLT_CD	분자병리NRASMUTATION 검사결과코드	N-ras mutation 결과코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_NREX_RSLT_NM	분자병리NRASMUTATION 검사결과명	N-ras mutation 결과명 예) detected
대장암_병리_분자	MLPT_BRME_RSLT_CD	분자병리BRAF_MUTATIO N검사결과코드	BRAF mutation 결과코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_BRME_RSLT_NM	분자병리BRAF_MUTATIO N검사결과명	BRAF mutation 결과명 예) detected
대장암_병리_분자	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_외과	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_외과	PT_SBST_NO	환자대체번호	개인고유번호 예) RN000000002
대장암_병리_외과	SGPT_ACPT_YMD	외과병리접수일자	외과병리접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_외과	SGPT_SEQ	외과병리순번	외과병리순번 (정수) 예) 1
대장암_병리_외과	SGPT_READ_YMD	외과병리판독일자	외과병리판독일자 YYYY-MM-DD

대장암_병리_외과	SGPT_CELL_DIFF_CD	외과병리세포분화코드	Cellular differentiation코드 01: Adenocarcinoma, WD 02: Adenocarcinoma, MD 03: Adenocarcinoma, PD 04: Mucinous ca 05: Carcinoid 99: Other
대장암_병리_외과	SGPT_CELL_DIFF_NM	외과병리세포분화명	Cellular differentiation명 예) Adenocarcinoma, WD
대장암_병리_외과	SGPT_PATL_T_STAG_VL	외과병리병리학적T병기값	외과병리병리학적T병기값 예) 0
대장암_병리_외과	SGPT_PATL_N_STAG_VL	외과병리병리학적N병기값	외과병리병리학적N병기값 예) 0
대장암_병리_외과	SGPT_PATL_M_STAG_VL	외과병리병리학적M병기값	외과병리병리학적M병기값 예) 0
대장암_병리_외과	SGPT_SRMV_LN_CNT	외과병리절제림프절수	절제한 전체 림프절 개수 예) 44
대장암_병리_외과	SGPT_MTST_LN_CNT	외과병리전이림프절수	절제한 림프절 중 전이된 림프절 개수 예) 1
대장암_병리_외과	SGPT_SRMG_PCTS_STA_T_CD	외과병리수술절제면근위 암조직크기상태코드	수술절제면에 암조직 잔존유무 상태 코드 (proximal) 1: involed 2: uninvoled 3: indeterminate 9: Other
대장암_병리_외과	SGPT_SRMG_PCTS_STA_T_NM	외과병리수술절제면근위 암조직크기상태명	수술절제면에 암조직 잔존유무 상태 명 (proximal) 예) involed
대장암_병리_외과	SGPT_SRMG_DCTS_STAT_CD	외과병리수술절제면원위 암조직크기상태코드	수술절제면에 암조직 잔존유무 상태 코드 (distal) 1: involed 2: uninvoled 3: indeterminate 9: Other
대장암_병리_외과	SGPT_SRMG_DCTS_STAT_NM	외과병리수술절제면원위 암조직크기상태명	수술절제면에 암조직 잔존유무 상태 명 (distal) 예) involed
대장암_병리_외과	SGPT_SRMG_RCTS_STAT_CD	외과병리수술절제면방사형암조직크기상태코드	수술절제면에 암조직 잔존유무 상태 코드 (radial(rectum only)) 1: involed 2: uninvoled 3: indeterminate 9: Other

대장암_병리_외과	SGPT_SRMG_RCTS_ST AT_NM	외과병리수술절제면방사 형암조직크기상태명	수술절제면에 암조직 잔존유무 상태 명 (radial(rectum only)) 예) involed
대장암_병리_외과	SGPT_NERV_PREX_CD	외과병리신경주위침윤코 드	신경주위침윤코드 1: present 2: absent 3: not identified 4: no record 9: Other
대장암_병리_외과	SGPT_NERV_PREX_NM	외과병리신경주위침윤명	신경주위침윤명 예) present
대장암_병리_외과	SGPT_VNIN_CD	외과병리정맥침윤코드	정맥침윤상태코드 1: present 2: absent 3: not identified 4: no record 9: Other
대장암_병리_외과	SGPT_VNIN_NM	외과병리정맥침윤명	정맥침윤상태명 예) present
대장암_병리_외과	SGPT_ANIN_CD	외과병리혈관조영침윤코 드	혈관조영침윤코드 1: present 2: absent 3: not identified 4: no record 9: Other
대장암_병리_외과	SGPT_ANIN_NM	외과병리혈관조영침윤명	혈관조영침윤명 예) present
대장암_병리_외과	SGPT_TUMR_BUDD_N M	외과병리종양발아명	Tumor budding 명 예) present
대장암_병리_외과	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_수술_정보	CENTER_CD	센터코드	센터코드 (2자리)
대장암_수술_정보	PT_SBST_NO	환자대체번호	개인고유번호 예) RN000000002
대장암_수술_정보	OPRT_YMD	수술일자	수술일자 (수집기준일자) YYYY-MM-DD
대장암_수술_정보	OPRT_SEQ	수술순번	수술순번 (정수) 예) 1

대장암_수술_정보	OPRT_CLCN_OPRT_KIN D_CD	수술대장암수술종류코드	대장암수술종류코드 01 : 결장절제술 02 : 반흔허니아근본수술 03 : 소장절제술 04 : 장 및 장간막 손상수술 05 : 장간막종양적출술 06 : 장관유착박리술 07 : 장루조성술 08 : 장루폐쇄술 09 : 장문합술 10 : 장폐색증수술 11 : 직장 및 에스장절제술 12 : 직장종양 절제술 13 : 직장항문 주위농양수술 14 : 진단적개복술 15 : 충수절제술 16 : 후복막종양적출술 99 : 기타
대장암_수술_정보	OPRT_CLCN_OPRT_KIN D_NM	수술대장암수술종류명	대장암수술종류코드명 예) 결장절제술
대장암_수술_정보	OPRT_CURA_RSCT_CD	수술근치적절제술코드	Curative resection 코드 1: 해당없음 2: curative 3: palliative 9: Other
대장암_수술_정보	OPRT_CURA_RSCT_NM	수술근치적절제술명	Curative resection명 예) curative
대장암_수술_정보	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_치료_항암제	CENTER_CD	센터코드	센터코드 (2자리)
대장암_치료_항암제	PT_SBST_NO	환자대체번호	개인고유번호 예) RN000000002
대장암_치료_항암제	CSTR_STRT_YMD	항암제치료시작일자	일자 YYYY-MM-DD
대장암_치료_항암제	CSTR_SEQ	항암제치료순번	항암제치료방순번 (정수) 예) 1

대장암_치료_항암제	CSTR_REGN_CD	항암제치료요법코드	Regimen 코드 01: none 02 :5-FU ± leucovorin 03 : capecitabine 04 : S-1 05 : UFT ± leucovorin 06 : avastin + fluoropyrimidine(FL/X/UFT/S-1) 07 : FOLFOX 08 : FOLFOX + AVASTIN 09 : FOLFOX + ERBITUX 10 : XELOX 11 : XELOX + AVASTIN 12 : XELOX + ERBITUX 13 :FOLFIRI 14 : FOLFIRI + AVASTIN 15 : FOLFIRI + ERBITUX 99 : Other
대장암_치료_항암제	CSTR_REGN_NM	항암제치료요법명	Regimen 명 예) S-1
대장암_치료_항암제	CSTR_NT	항암제치료횟수	항암치료횟수(Line) 예) 10
대장암_치료_항암제	CSTR_END_YMD	항암제치료종료일자	일자 YYYY-MM-DD
대장암_치료_항암제	CSTR_PRPS_CD	항암제치료목적코드	항암치료목적코드 01: Neo-adjuvant 02: Adjuvant 03: Palliative 04: Concurrent 05: Induction 06: Maintenance 07: Salvage 08: Consolidation 99: Other
대장암_치료_항암제	CSTR_PRPS_NM	항암제치료목적명	항암치료목적명 예) Adjuvant
대장암_치료_항암제	CSTR_CYCL_VL	항암제치료주기값	항암치료주기값 예) 5
대장암_치료_항암제	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_치료_방사선	CENTER_CD	센터코드	센터코드 (2자리)
대장암_치료_방사선	PT_SBST_NO	환자대체번호	개인고유번호 예) RN000000002
대장암_치료_방사선	RDT_STRT_YMD	방사선치료시작일자	방사선치료시작일자 (수집기준일자) YYYY-MM-DD

대장암_치료_방사선	RDT_RT_NO	방사선치료RT번호	방사선치료RT번호 예) 1
대장암_치료_방사선	RDT_SITE_CD	방사선치료부위코드	치료부위코드 예) 0851 ※ 치료부위명 참조
대장암_치료_방사선	RDT_SITE_NM	방사선치료부위명	방사선치료부위명 예) Pelvis Whole
대장암_치료_방사선	RDT_END_YMD	방사선치료종료일자	일자 YYYY-MM-DD
대장암_치료_방사선	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_사망_정보	CENTER_CD	센터코드	센터코드 (2자리)
대장암_사망_정보	PT_SBST_NO	환자대체번호	개인고유번호 예) RN000000002
대장암_사망_정보	DEAD_YMD	사망일자	사망일자 (수집기준일자) YYYY-MM-DD
대장암_사망_정보	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_진단_재발정 보	CENTER_CD	센터코드	센터코드 (2자리)
대장암_진단_재발정 보	PT_SBST_NO	환자대체번호	개인고유번호 예) RN000000002
대장암_진단_재발정 보	RLPS_DIAG_YMD	재발진단일자	환자의 재발진단일 (수집기준일자) YYYY-MM-DD
대장암_진단_재발정 보	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()

- 국립암센터

중분류 (Table)	변수명 (Column)	변수 설명	변수값(코드) 설명 (99=모름/기타/미측정)
대장암_환자_기본정보	CENTER_CD	센터코드	센터코드 (2자리)
대장암_환자_기본정보	PT_SBST_NO	환자대체번호	개인고유번호 예) CL000000002
대장암_환자_기본정보	BSPT_SEX_CD	기본환자성별코드	환자의 성별 M : 남자 F : 여자 E : 기타
대장암_환자_기본정보	BSPT_BRYM	기본환자생년월	생년월 YYYYMM
대장암_환자_기본정보	BSPT_FRST_DIAG_YMD	기본환자최초진단일자	대장암초진일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_FRST_DIAG_CD	기본환자최초진단코드	암진단코드 예) C18
대장암_환자_기본정보	BSPT_FRST_DIAG_NM	기본환자최초진단명	암진단명 예) colon
대장암_환자_기본정보	BSPT_IDGN_AGE	기본환자진단시연령	대장암 최초 진단시 나이(만) 예) 45 ※주진단으로 C18-20(하위코드포함)를 받은 환자
대장암_환자_기본정보	BSPT_FRST_OPRT_YMD	기본환자최초수술일자	최초 대장암 진단 이후 첫수술일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_FRST_ANCN_TRT_M_STRT_YMD	기본환자최초항암치료시작일자	최초항암치료시작일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_FRST_RDT_STRT_YMD	기본환자최초방사선치료시작일자	최초 대장암 진단 이후 첫 방사선치료시작일자 YYYY-MM-DD
대장암_환자_기본정보	BSPT_DEAD_YMD	기본환자사망일자	사망일자 YYYY-MM-DD
대장암_환자_기본정보	CENTER_LAST_VST_YMD	센터최종방문일자	센터 최종 방문일 YYYY-MM-DD
대장암_환자_기본정보	OVRL_SRLV_DTRN_DCNT	전체생존기간일수	환자의 진단일부터 사망일 또는 최종내원일까지의 기간(일)
대장암_환자_기본정보	BSPT_STAG_VL	기본환자병기값	기본환자병기값 예) 0
대장암_환자_기본정보	BSPT_T_STAG_VL	기본환자T병기값	기본환자T병기값 예) 0
대장암_환자_기본정보	BSPT_N_STAG_VL	기본환자N병기값	기본환자N병기값 예) 0

대장암_환자_기본정보	BSPT_M_STAG_VL	기본환자M병기값	기본환자M병기값 예) 0
대장암_환자_기본정보	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_검사_진단	CENTER_CD	센터코드	센터코드 (2자리)
대장암_검사_진단	PT_SBST_NO	환자대체번호	개인고유번호 예) CL000000002
대장암_검사_진단	CEXM_YMD	진단검사일자	진단검사일자 (수집기준일자) YYYY-MM-DD
대장암_검사_진단	CEXM_SEQ	진단검사순번	진단검사순번 (정수) 예) 1
대장암_검사_진단	CEXM_NM	진단검사명	진단검사명 예) CEA
대장암_검사_진단	CEXM_RSLT_CONT	진단검사결과내용	진단검사 결과 내용 예) 33.2
대장암_검사_진단	CEXM_RSLT_UNIT_CONNT	진단검사결과단위내용	진단검사 결과 내용의 단위 예) ng/ml
대장암_검사_진단	CEXM_RSLT_REF_CONNT	진단검사결과참고내용	진단검사 종류 참고치 예) 0~5.0
대장암_검사_진단	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_생체검사	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_생체검사	PT_SBST_NO	환자대체번호	개인고유번호 예) CL000000002
대장암_병리_생체검사	BPTH_ACPT_YMD	생체검사병리접수일자	생검 접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_생체검사	BPTH_SEQ	생체검사병리순번	생검 순번 (정수) 예) 1
대장암_병리_생체검사	BPTH_READ_YMD	생체검사병리판독일자	생검 판독일자 YYYY-MM-DD
대장암_병리_생체검사	BPTH_SITE_CONT	생체검사병리부위내용	시행된 생체검사 부위 예) colon
대장암_병리_생체검사	BPTH_BPSY_RSLT_CONNT	생체검사병리생체검사결과내용	생체검사 결과 내용 예) carcinoma
대장암_병리_생체검사	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_면역	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_면역	PT_SBST_NO	환자대체번호	개인고유번호 예) CL000000002
대장암_병리_면역	IMPT_ACPT_YMD	면역병리접수일자	면역병리검사접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_면역	IMPT_SEQ	면역병리순번	면역병리 정수 순번 예) 1
대장암_병리_면역	IMPT_READ_YMD	면역병리판독일자	면역병리판독일자 YYYY-MM-DD

대장암_병리_면역	IMPT_HM1E_RSLT_CD	면역병리HMLH1검사결과 코드	hMLH1 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HM1E_RSLT_NM	면역병리HMLH1검사결과 명	hMLH1 결과 명 예) loss
대장암_병리_면역	IMPT_HS2E_RSLT_CD	면역병리HMSH2검사결과 코드	hMLH2 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HS2E_RSLT_NM	면역병리HMSH2검사결과 명	hMLH2 결과 명 예) loss
대장암_병리_면역	IMPT_HS6E_RSLT_CD	면역병리HMSH6검사결과 코드	hMLH6 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HS6E_RSLT_NM	면역병리HMSH6검사결과 명	HMSH6 결과 명 예) loss
대장암_병리_면역	IMPT_HP2E_RSLT_CD	면역병리HPMS2검사결과 코드	hPMS2 결과 코드 1: loss 2: intact 9: Other
대장암_병리_면역	IMPT_HP2E_RSLT_NM	면역병리HPMS2검사결과 명	HPMS2 결과 명 예) loss
대장암_병리_면역	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_분자	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_분자	PT_SBST_NO	환자대체번호	개인고유번호 예) CL000000002
대장암_병리_분자	MLPT_ACPT_YMD	분자병리접수일자	분자병리검사접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_분자	MLPT_SEQ	분자병리순번	분자병리순번 예) 1
대장암_병리_분자	MLPT_READ_YMD	분자병리판독일자	면역병리판독일자 YYYY-MM-DD
대장암_병리_분자	MLPT_MSIE_RSLT_CD	분자병리MSI검사결과코드	MSI검사결과코드 1 : stable 2 : low 3 : high 9 : Other
대장암_병리_분자	MLPT_MSIE_RSLT_NM	분자병리MSI검사결과명	MSI검사결과명 예) stable

대장암_병리_분자	MLPT_KE2E_RSLT_CD	분자병리KRAS_MUTATION_EXON2검사결과코드	K-ras mutation (exon2) 결과 코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_KE2E_RSLT_NM	분자병리KRAS_MUTATION_EXON2검사결과명	K-ras mutation (exon2) 결과명 예) not detected
대장암_병리_분자	MLPT_KRES_RSLT_CD	분자병리KRAS_MUTATION 검사결과코드	K-ras mutation 결과 코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_KRES_RSLT_NM	분자병리KRAS_MUTATION 검사결과명	K-ras mutation 결과명 예) not detected
대장암_병리_분자	MLPT_NREX_RSLT_CD	분자병리NRAS_MUTATION 검사결과코드	N-ras mutation 결과코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_NREX_RSLT_NM	분자병리NRAS_MUTATION 검사결과명	N-ras mutation 결과명 예) detected
대장암_병리_분자	MLPT_BRME_RSLT_CD	분자병리BRAF_MUTATION 검사결과코드	BRAF mutation 결과코드 1: not detected 2: detected 9: Other
대장암_병리_분자	MLPT_BRME_RSLT_NM	분자병리BRAF_MUTATION 검사결과명	BRAF mutation 결과명 예) detected
대장암_병리_분자	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_병리_외과	CENTER_CD	센터코드	센터코드 (2자리)
대장암_병리_외과	PT_SBST_NO	환자대체번호	개인고유번호 예) CL000000002
대장암_병리_외과	SGPT_ACPT_YMD	외과병리접수일자	외과병리접수일자 (수집기준일자) YYYY-MM-DD
대장암_병리_외과	SGPT_SEQ	외과병리순번	외과병리순번 (정수) 예) 1
대장암_병리_외과	SGPT_READ_YMD	외과병리판독일자	외과병리판독일자 YYYY-MM-DD

대장암_병리_외과	SGPT_CELL_DIFF_CD	외과병리세포분화코드	Cellular differentiation코드 01: Adenocarcinoma, WD 02: Adenocarcinoma, MD 03: Adenocarcinoma, PD 04: Mucinous ca 05: Carcinoid 99: Other
대장암_병리_외과	SGPT_CELL_DIFF_NM	외과병리세포분화코드	Cellular differentiation코드 01: Adenocarcinoma, WD 02: Adenocarcinoma, MD 03: Adenocarcinoma, PD 04: Mucinous ca 05: Carcinoid 99: Other
대장암_병리_외과	SGPT_PATL_STAG_VL	외과병리병리학적병기값	외과병리병리학적병기값 예) pT3N2bM1a
대장암_병리_외과	SGPT_PATL_T_STAG_VL	외과병리병리학적T병기값	외과병리병리학적T병기값 예) 0
대장암_병리_외과	SGPT_PATL_N_STAG_VL	외과병리병리학적N병기값	외과병리병리학적N병기값 예) 0
대장암_병리_외과	SGPT_PATL_M_STAG_VL	외과병리병리학적M병기값	외과병리병리학적M병기값 예) 0
대장암_병리_외과	SGPT_SRMV_LN_CNT	외과병리절제림프절수	절제한 전체 림프절 개수 예) 44
대장암_병리_외과	SGPT_MTST_LN_CNT	외과병리전이림프절수	절제한 림프절 중 전이된 림프절 개수 예) 1
대장암_병리_외과	SGPT_NERV_PREX_CD	외과병리신경주위침윤코드	신경주위침윤코드 1: present 2: absent 3: not identified 4: no record 9: Other
대장암_병리_외과	SGPT_NERV_PREX_NM	외과병리신경주위침윤명	신경주위침윤명 예) present
대장암_병리_외과	SGPT_VNIN_CD	외과병리정맥침윤코드	정맥침윤상태코드 1: present 2: absent 3: not identified 4: no record 9: Other
대장암_병리_외과	SGPT_VNIN_NM	외과병리정맥침윤명	정맥침윤상태명 예) present
대장암_병리_외과	SGPT_ANIN_CD	외과병리혈관조영침윤코드	혈관조영침윤코드 1: present 2: absent 3: not identified 4: no record 9: Other
대장암_병리_외과	SGPT_ANIN_NM	외과병리혈관조영침윤명	혈관조영침윤명 예) present
대장암_병리_외과	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()

대장암_수술_정보	CENTER_CD	센터코드	센터코드 (2자리)
대장암_수술_정보	PT_SBST_NO	환자대체번호	개인고유번호 예) CL000000002
대장암_수술_정보	OPRT_YMD	수술일자	수술일자 (수집기준일자) YYYY-MM-DD
대장암_수술_정보	OPRT_SEQ	수술순번	수술순번 (정수) 예) 1
대장암_수술_정보	OPRT_CLCN_OPRT_KIN_D_CD	수술대장암수술종류코드	대장암수술종류코드 01 : 결장절제술 02 : 반흔허니아근본수술 03 : 소장절제술 04 : 장 및 장간막 손상수술 05 : 장간막중앙적출술 06 : 장관유착박리술 07 : 장루조성술 08 : 장루폐쇄술 09 : 장문합술 10 : 장폐색증수술 11 : 직장 및 에스장절제술 12 : 직장중앙 절제술 13 : 직장항문 주위농양수술 14 : 진단적개복술 15 : 충수절제술 16 : 후복막중앙적출술 99 : 기타
대장암_수술_정보	OPRT_CLCN_OPRT_KIN_D_NM	수술대장암수술종류명	대장암수술종류코드명 예) 결장절제술
대장암_수술_정보	OPRT_CURA_RSCT_CD	수술근치적절제술코드	Curative resection 코드 1: 해당없음 2: curative 3: palliative 9: Other
대장암_수술_정보	OPRT_CURA_RSCT_NM	수술근치적절제술명	Curative resection명 예) curative
대장암_수술_정보	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_치료_항암제	CENTER_CD	센터코드	센터코드 (2자리)
대장암_치료_항암제	PT_SBST_NO	환자대체번호	개인고유번호 예) CL000000002
대장암_치료_항암제	CSTR_STRT_YMD	항암제치료시작일자	일자 YYYY-MM-DD
대장암_치료_항암제	CSTR_SEQ	항암제치료순번	항암제처방순번 (정수) 예) 1

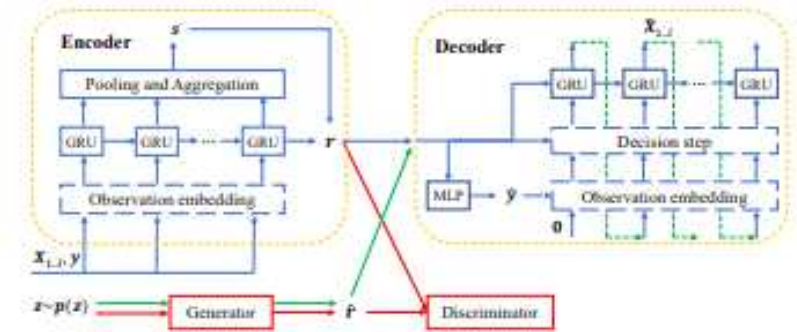
대장암_치료_항암제	CSTR_REGN_CD	항암제치료요법코드	Regimen 코드 01: none 02 :5-FU ± leucovorin 03 : capecitabine 04 : S-1 05 : UFT ± leucovorin 06 : avastin + fluoropyrimidine(FL/X/UFT/S-1) 07 : FOLFOX 08 : FOLFOX + AVASTIN 09 : FOLFOX + ERBITUX 10 : XELOX 11 : XELOX + AVASTIN 12 : XELOX + ERBITUX 13 :FOLFIRI 14 : FOLFIRI + AVASTIN 15 : FOLFIRI + ERBITUX 99 : Other
대장암_치료_항암제	CSTR_REGN_NM	항암제치료요법명	Regimen 명 예) S-1
대장암_치료_항암제	CSTR_NT	항암제치료횟수	항암치료횟수(Line) 예) 10
대장암_치료_항암제	CSTR_END_YMD	항암제치료종료일자	일자 YYYY-MM-DD
대장암_치료_항암제	CSTR_PRPS_CD	항암제치료목적코드	항암치료목적코드 01: Neo-adjuvant 02: Adjuvant 03: Palliative 04: Concurrent 05: Induction 06: Maintenance 07: Salvage 08: Consolidation 99: Other
대장암_치료_항암제	CSTR_PRPS_NM	항암제치료목적명	항암치료목적명 예) Adjuvant
대장암_치료_항암제	CSTR_CYCL_VL	항암제치료주기값	항암치료주기값 예) 5
대장암_치료_항암제	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_치료_방사선	CENTER_CD	센터코드	센터코드 (2자리)
대장암_치료_방사선	PT_SBST_NO	환자대체번호	개인고유번호 예) CL000000002
대장암_치료_방사선	RDT_STRT_YMD	방사선치료시작일자	방사선치료시작일자 (수집기준일자) YYYY-MM-DD

대장암_치료_방사선	RDT_RT_NO	방사선치료RT번호	방사선치료RT번호 예) 1
대장암_치료_방사선	RDT_END_YMD	방사선치료종료일자	일자 YYYY-MM-DD
대장암_치료_방사선	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()
대장암_사망_정보	CENTER_CD	센터코드	센터코드 (2자리)
대장암_사망_정보	PT_SBST_NO	환자대체번호	개인고유번호 예) CL000000002
대장암_사망_정보	DEAD_YMD	사망일자	사망일자 (수집기준일자) YYYY-MM-DD
대장암_사망_정보	CRTN_DT	생성일시	생성일시 DEFAULT current_timestamp()

□ 데이터 구축 과정

○ GAN 기반 합성데이터 생성

- RTSGAN은 데이터의 생성주기가 불규칙한 의료 데이터 합성에 특화된 모델로, Encoder-Decoder 모듈과 Generation 모듈로 구성되어 있음.
- Encoder-Decoder 모듈에서는 데이터를 고정된 차원의 latent space로 feature를 추출하고, Generation 모듈에서는 추출된 feature를 합성함.
- 학습 과정은 다음과 같음, 먼저 의료 데이터의 경우, 시간이 지남에 따라 값이 변하는 동적 변수와 시간과 상관없이 변하지 않는 정적 변수로 구성되어 있으며, 증상에 따라 검사가 이루어지기 때문에 변수 별 결측 비율이 매우 높다는 특징이 있음. 이는 모델이 데이터의 정보를 학습하는데 어려움을 유발함. 때문에 모델에서는 먼저 동적 변수와 정적 변수 뿐만 아니라 이전 시점의 관측값과 언제 관측되었는지에 대한 시간 정보를 고려하여 임베딩을 진행함. 이후 Encoder의 stacked RNN을 거쳐 Fully connected layer를 통해 고정된 길이의 latent vector를 구함. 다음으로 Decoder에서는 먼저 정적변수를 복원한 후 이를 활용하여 동적 변수를 복원함. 이때 어떤 데이터가 어느 시점에 관측이 될 것인지를 먼저 결정한 후 해당 정보를 바탕으로 어떤 값을 가질 것인지를 복원함.
- Generation 모듈에서는 Encoder-Decoder 모듈을 통해 구해진 latent vector z 를 GAN(WGAN-GP)을 통해 합성함.
- 마지막으로 Generation 모듈을 통해 생성된 \hat{z} 를 Decoder의 입력으로 사용하여 합성 데이터를 생성함.



[그림 1] RTSGAN 구조

□ 데이터 품질 검증

○ GAN 기반 합성데이터 검증

- 환자 기본정보
 - 최초 진단일자보다 빠른 치료시작일자 기록 삭제
 - 기본환자 병기가 3기일 때, T 병기와 N 병기가 0인 기록 삭제
- 병리 외과
 - 외과 병리 전체 림프절 수 데이터가 있으나 외과 병리 전이 림프절 수 데이터가 없는 기록 삭제
 - 외과 병리 전이 림프절 수가 1~3인 데이터에 대해 N 병기 값이 1이 아닌 기록에 대해 1로 일괄 변경
- 항암제 치료
 - 사망일자보다 늦은 항암제 치료 종료일자 기록 삭제
- 방사선 치료
 - 사망일자보다 늦은 방사선 치료 종료일자 기록 삭제

□ GAN 기반 합성 데이터 관리 담당자

- 수행기관 : 국립암센터
 - 담당자 : 류광선 선임연구원
 - 전 화 : 031-920-0652
 - 이메일 : niceplay13@ncc.re.kr

□ 사사

- This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number : HI21C0974).
- Synthetic data produced by utilizing the colorectal cancer clinical library data built on the healthcare big data platform (CONNECT) supported by the national information society agency(NIA)