EMBRACE-II Variable Description

Version Date: October 23, 2025
Author: Stefan Ecker, PhD

Quick Reference: All Added Variables EMBRACE-II

This document describes the custom variables added to the EMBRACE-II dataset during data processing. Variables are organized by functional categories.

1. Patient and Study Information

Variable	Details
`age`	• Patient age at diagnosis • **Calculation:** `histology_assessment_date` − `year_of_birth` • **Implementation:** `add_age()` • **Unit:** Years
`included_in_study`	• Indicates whether a patient is included in the final study cohort • **Calculation:** Based on the study's inclusion/exclusion criteria documented in the official study flowchart • **Implementation:** `load_embrace_ii()` • **Values:** TRUE = included in study, FALSE = excluded from study
`withdrew_consent`	• Indicates if a patient withdrew consent from the study • **Calculation:** Based on `followup_*m` columns • **Implementation:** `add_lost_to_fu()` • **Values:** TRUE = consent withdrawn, FALSE = consent maintained
`is_lost_to_fu`	• Indicates if a patient was lost to follow-up during the study • **Calculation:** Based on `followup_*m` columns • **Implementation:** `add_lost_to_fu()` • **Values:** TRUE = patient lost to followup, FALSE = patient not lost to followup

2. Diagnostic Measurements

Variable	Details
`max_tumor_dimension_ sta_d`	• Maximum tumor dimension at diagnosis • **Calculation:** Overall maximum value across both MRI and gynecological exam measurements (`mri_tumor_*_sta_d` and `gyn_tumor_*_sta_d`) • **Implementation:** `add_max_tumor_dimension()` • **Unit:** Millimeters
`gyn_max_parametrium_ sta_d`	• Maximum parametrial involvement from gynecological examination at diagnosis • **Calculation:** Maximum value between left and right parametrial involvement (`gyn_left_parametrium_sta_d` and `gyn_right_parametrium_sta_d`) • **Implementation:** `add_parametrial_involvement()` • **Values:** 0 = Not involved, 1 = Proximal, 2 = Distal, 3 = To pelvic wall
`gyn_max_tumor_ dimension_sta_d`	• Maximum tumor dimension from gynecological examination • **Calculation:** Maximum value across all gynecological tumor measurements (width, height, thickness) • **Implementation:** `add_max_tumor_dimension()` • **Unit:** Millimeters
`mri_max_parametrium_ sta_d`	• Maximum parametrial involvement from MRI at diagnosis • **Calculation:** Maximum value between left and right parametrial involvement (`mri_left_parametrium_sta_d` and `mri_right_parametrium_sta_d`) • **Implementation:** `add_parametrial_involvement()` • **Values:** 0 = Not involved, 1 = Proximal, 2 = Distal, 3 = To pelvic wall
`mri_max_tumor_ dimension_sta_d`	• Maximum tumor dimension from MRI • **Calculation:** Maximum value across all MRI tumor measurements (width, height, thickness) • **Implementation:** `add_max_tumor_dimension()` • **Unit:** Millimeters

3. Brachytherapy Treatment and Technique

Variable	Details
`icis`	• Intracavitary/Interstitial classification • **Calculation:** Based on `fraction_*_technique_tdvh` columns • **Implementation:** `add_icis()` • **Values:** Binary (1 = ICIS, 0 = IC) • **Note:** Derived from technique values across fractions
`icis_parallel_oblique`	• Detailed ICIS needle classification • **Calculation:** Based on `fraction_*_technique_tdvh` columns • **Implementation:** `add_parallel_oblique_needles()` • **Values:** "no needles", "parallel needles", "oblique needles", "other"
`average_nr_active_ needles`	• Average number of active needles • **Calculation:** Mean of `fraction_*_active_needles` columns • **Implementation:** `add_average_active_needles()` • **Unit:** Number of needles
`time_to_bt`	• Time from EBRT start to first BT fraction • **Calculation:** `fraction01date_tdvh` - `ebrt_start_date_tdvh` • **Implementation:** `add_time_to_bt()` • **Unit:** Days
`time_to_bt_percent`	• Relative timing of brachytherapy • **Calculation:** `time_to_bt` / `ott_ebrt` • **Implementation:** `add_time_to_bt()` • **Unit:** Proportion
`ott_ebrt`	• Overall treatment time of EBRT • **Calculation:** `ebrt_end_date_tdvh` - `ebrt_start_date_tdvh` • **Implementation:** `add_time_to_bt()` • **Unit:** Days
`ott`	• Overall Treatment Time from start of EBRT to last treatment • **Calculation:** Time interval between `ebrt_start_date_tdvh` and `last_treatment_date` (which is the latest of EBRT end date and all fraction dates) • **Implementation:** `emii_add_ott()` • **Unit:** Days
`last_treatment_date`	• Date of the last treatment received • **Calculation:** Maximum date among EBRT end date (`ebrt_end_date_tdvh`) and all brachytherapy fraction dates (`fraction*date_tdvh`) • **Implementation:** `emii_add_ott()` • **Unit:** Date

Variable	Details
`fraction01hrctv_ volume_bins`	• CTV-HR volume categories • **Calculation:** Based on `fraction01hrctv_volume_tdvh` • **Implementation:** `add_hrctv_volume_bins()` • **Values:** 0: < 30 cm^3, 1: 30-45 cm^3, 2: > 45 cm^3, NA: missing
`trak_total_sum`	• Total Reference Air Kerma (TRAK) summed across all brachytherapy fractions • **Calculation:** Sum of all `fraction*trak_tdvh` values • **Implementation:** `emii_add_trak_absolute()` • **Unit:** cGy at 1m
`trak_needles_sum`	• Total TRAK contribution from needles across all brachytherapy fractions • **Calculation:** Sum of `fraction*_needles_abs` values, which are derived from percentage contributions (`trak * trak_needles_pct / 100`) • **Implementation:** `emii_add_trak_absolute()` • **Unit:** cGy at 1m
`trak_vaginal_ applicator_sum`	• Total TRAK contribution from vaginal applicator across all fractions • **Calculation:** Sum of `fraction*_vaginal_applicator_abs` values, which are derived from percentage contributions (`trak * trak_vaginal_applicator_pct / 100`) • **Implementation:** `emii_add_trak_absolute()` • **Unit:** cGy at 1m
`trak_tandem_ applicator_sum`	• Total TRAK contribution from tandem applicator across all fractions • **Calculation:** Sum of `fraction*_tandem_applicator_abs` values, which are derived from percentage contributions (`trak * trak_tandem_applicator_pct / 100`) • **Implementation:** `emii_add_trak_absolute()` • **Unit:** cGy at 1m
`imp1gyn_max_ parametrium_sta_b`	• Maximum parametrial involvement from gynecological examination at first implant • **Calculation:** Maximum value between left and right parametrial involvement (`imp1gyn_left_parametrium_sta_b` and `imp1gyn_right_parametrium_sta_b`) • **Implementation:** `add_parametrial_involvement()` • **Values:** 0 = Not involved, 1 = Proximal, 2 = Distal, 3 = To pelvic wall

Variable	Details
`imp1image_max_ parametrium_sta_b`	• Maximum parametrial involvement from imaging at first implant • **Calculation:** Maximum value between left and right parametrial involvement (`imp1image_left_parametrium_sta_b` and `imp1imaging_right_parametrium_sta_b`) • **Implementation:** `add_parametrial_involvement()` • **Values:** 0 = Not involved, 1 = Proximal, 2 = Distal, 3 = To pelvic wall

4. External Beam Radiotherapy (EBRT)

Variable	Details
`ebrt_elective_target_ algorithm`	• Algorithm-recommended treatment target volume for elective nodal irradiation • **Calculation:** Determined based on EMBRACE-II protocol • **Implementation:** `emii_add_elective_targets()` • **Values:** 'Small Pelvis', 'Large Pelvis', 'Large Pelvis + Para-aortic', 'Large Pelvis + Inguinal', 'Large Pelvis + Para-aortic + Inguinal'
`elective_high_risk`	• Indicates if high-risk criteria for nodal treatment are met • **Calculation:** TRUE if pathological nodes present AND >= 3 nodes found, OR if any common iliac or para-aortic nodes present • **Implementation:** `emii_add_elective_targets()` • **Values:** TRUE/FALSE
`elective_inguinal`	• Indicates if inguinal node irradiation criteria are met • **Calculation:** TRUE if maximum vaginal involvement = 3 (distal third) OR any inguinal nodes present in right or left groin • **Implementation:** `emii_add_elective_targets()` • **Values:** TRUE/FALSE
`elective_low_risk`	• Indicates if low-risk criteria for nodal treatment are met • **Calculation:** TRUE if NO pathological nodes, tumor <= 40mm, TNM stage 1/2/3/5, squamous cell histology, and no corpus uteri involvement on MRI • **Implementation:** `emii_add_elective_targets()` • **Values:** TRUE/FALSE
`any_node_ci_pa`	• Indicates presence of any common iliac or para-aortic lymph nodes at diagnosis • **Calculation:** TRUE if any node has anatomical position of "R com iliac", "L com iliac", or "Para Aortic" • **Implementation:** `emii_add_elective_targets()` • **Values:** TRUE/FALSE

5. Lymph Nodes and Metastases

Variable	Details
`nodal_classification`	• Classification of nodal involvement status • **Calculation:** Based on pathological nodes present (`pathological_nodes_present`) and Para-Aortic lymph node involvement at diagnosis • **Implementation:** `emii_add_nodal_classification()` • **Values:** 'N0' (no pathological nodes), 'N1PAN' (pathological nodes with Para-Aortic involvement), 'N1pelvic' (pathological nodes without Para-Aortic involvement)
`number_common_iliac_ln_ stat_d`	• Count of common iliac lymph nodes at diag. • **Calculation:** Sum of left and right common iliac lymph nodes from staging data • **Implementation:** `emii_add_number_common_iliac_ln_stat_d()` • **Unit:** Count
`number_paraaortic_ln_ stat_d`	• Count of para-aortic lymph nodes at diag. • **Calculation:** Count of lymph nodes with "Para Aortic" anatomical position from staging data • **Implementation:** `add_number_paraaortic_ln_stat_d()` • **Unit:** Count
`has_*_followup`	• Indicates lymph node involvement during follow-up by specific anatomical location • **Calculation:** Checks if any lymph node documented during follow-up has this specific anatomical position • **Implementation:** `emii_add_recurrent_nodes()` • **Example variables:** `has_L.ext.iliac_followup`, `has_Para.Aortic_followup`, etc. • **Values:** TRUE/FALSE
`has_*_diagnosis`	• Indicates lymph node involvement at diagnosis by specific anatomical location • **Calculation:** Checks if any lymph node documented at diagnosis has this specific anatomical position • **Implementation:** `emii_add_diagnostic_nodes()` • **Example variables:** `has_L.ext.iliac_diagnosis`, `has_Para.Aortic_diagnosis`, etc. • **Values:** TRUE/FALSE
`has_*_metastases`	• Indicates presence of specific metastatic disease sites • **Calculation:** Checks if any follow-up visit reports metastasis at this site • **Implementation:** `add_metastases()` • **Example variables:** `has_liver_metastases`, `has_bone_metastases`, `has_brain_metastases`, `has_lung_metastases`, `has_other_metastases` • **Values:** TRUE/FALSE

Variable	Details
`has_*_nodes`	• Indicates presence of distant nodal metastases • **Calculation:** Checks if any follow-up visit reports nodes at these sites • **Implementation:** `add_metastases()` • **Example variables:** `has_paraaortic_nodes_above_l2`, `has_supraclavicular_nodes`, `has_mediastinal_nodes` • **Values:** TRUE/FALSE
`has_abdominal_ carcinomatosis`	• Indicates presence of abdominal carcinomatosis • **Calculation:** Checks if any follow-up visit reports abdominal carcinomatosis • **Implementation:** `add_metastases()` • **Values:** TRUE/FALSE

6. Event Endpoints

Variable	Endpoint	Details
`event_localfailure`	Local Control	• **Calculation:** Triggered when any followup `disease_local_status_*` column has a value of 2, indicating recurrence/progression • **Implementation:** `add_local_failure_event()` • **Values:** 1 = failure (event occurred), 0 = no failure (no event)
`event_nodalfailure`	Nodal Control (lower PAO)	• **Calculation:** Triggered when any followup `disease_nodal_status_*` column has a value of 2, indicating nodal recurrence or progression • **Implementation:** `add_nodal_failure_event()` • **Values:** 1 = failure (event occurred), 0 = no failure (no event)
`event_nodalcontrol_incl_pao`	Nodal control (incl. all PAO)	• **Calculation:** Triggered when either pelvic nodal failure (`event_pelvic_nodal`) OR para-aortic nodal failure (`event_paraaortic_nodal`) occurs • **Implementation:** `emii_add_nodalcontrol_incl_pao()` • **Values:** 1 = failure (event occurred), 0 = no failure (no event)
`event_systemicfailure`	Systemic Control (incl. upper PAO)	• **Calculation:** Triggered when any followup `disease_systemic_status_*` column has a value of 2, indicating systemic recurrence or progression • **Implementation:** `add_systemic_failure_event()` • **Values:** 1 = failure (event occurred), 0 = no failure (no event)
`event_systemic_excl_pao`	Systemic control (excl. PAO)	• **Calculation:** All systemic failures | (`event_systemicfailure`) MINUS cases where | para-aortic nodes above L2 are the ONLY | systemic failure. Cases with para-aortic nodes above L2 plus any other systemic | metastasis remain included • **Implementation:** | `emii_add_systemic_excl_pao()` | • **Values:** 1 = failure (event occurred), | 0 = no failure (no event) |

Variable	Endpoint	Details
`event_vitalstatus`	Overall Survival	• **Calculation:** Uses the value from `vital_status` when available; otherwise sets to 0 (alive) if a disease assessment date exists (`latest_assessment_date_disease`) • **Implementation:** `add_vitalstatus_event()` • **Values:** 1 = death (event occurred), 0 = alive (no event)
`event_pelvic_nodal`	Pelvic Nodal control	• **Calculation:** Triggered when nodal failure (`event_nodalfailure`) occurs AND at least one node is positive in any pelvic location: external iliac, internal iliac, common iliac, or parametrial/paracervical regions (left or right) • **Implementation:** `emii_add_pelvic_nodal_event()` • **Values:** 1 = failure (event occurred), 0 = no failure (no event)
`event_pelvic`	Pelvic control (local + nodal)	• **Calculation:** Triggered when either local failure (`event_localfailure`) OR pelvic nodal failure (`event_pelvic_nodal`) occurs • **Implementation:** `emii_add_pelvic_event()` • **Values:** 1 = failure (event occurred), 0 = no failure (no event)
`event_paraaortic_nodal`	Para-aortic nodal control	• **Calculation:** Triggered when either para-aortic nodes above L2 (`has_paraaortic_nodes_above_l2`) OR para-aortic nodes below L2 (`has_Para.Aortic_followup`) is detected • **Implementation:** `emii_add_paraaortic_nodal()` • **Values:** 1 = failure (event occurred), 0 = no failure (no event)
`event_locoregional`	Locoregional (Pelvic/PAO)	• **Calculation:** Triggered when either local failure (`event_localfailure`) OR nodal control including para-aortic (`event_nodalcontrol_incl_pao`) occurs • **Implementation:** `add_locoregional_event()` • **Values:** 1 = failure (event occurred), 0 = no failure (no event)

Variable	Endpoint	Details
`event_locoregional_alone`	Locoregional alone (no metastases)	• **Calculation:** Triggered when a locoregional failure occurs (`event_locoregional=1`) WITHOUT any concurrent distant metastases (`event_systemic_excl_pao=0`) • **Implementation:** `add_locoregional_event()` • **Values:** 1 = isolated locoregional recurrence, 0 = no recurrence or recurrence with distant metastases
`event_cancer_specific`	Cancer-Specific Survival	• **Calculation:** Triggered when death occurs (`event_vitalstatus=1`) AND the main cause of death is primary cancer (codes 1 or 3 in `vital_status_cause_of_death_vital_status`) • **Implementation:** `emii_add_cancer_specific()` • **Values:** 1 = cancer death, 0 = alive or non-cancer death
`event_disease_control`	Disease Control	• **Calculation:** Triggered when any failure occurs at any site: local failure OR nodal failure OR systemic failure • **Implementation:** `emii_add_disease_control()` • **Values:** 1 = any disease failure, 0 = disease controlled at all sites
`event_progression_free`	Progression-Free Survival	• **Calculation:** Triggered when any disease event (`event_disease_control=1`) OR death occurs (`event_vitalstatus=1`) • **Implementation:** `emii_add_progression_free_survival()` • **Values:** 1 = progression or death, 0 = alive without progression
`event_distant_alone`	Distant Metastases Alone	• **Calculation:** Triggered when systemic failure excluding para-aortic nodes (`event_systemic_excl_pao=1`) occurs AND | no locoregional failure (`event_locoregional=0`) | • **Implementation:** `emii_add_distant_alone()` • **Values:** 1 = isolated distant metastases, 0 = no distant metastases or concurrent locoregional disease

Variable	Endpoint	Details
`timetoevent_disease`	Time to Disease Event	• **Calculation:** Time interval in months from histology assessment date to either the first disease event date or the latest assessment date if no event occurred • **Implementation:** `add_time_to_diseaseevent()` • **Values:** Number of months
`timetoevent_vitalstatus`	Time to Last Vital Status	• **Calculation:** Time interval in months from histology assessment date to the latest vital status date • **Implementation:** `add_time_to_last_vitalstatus()` • **Values:** Number of months
`latest_vital_status_ date`	Latest Vital Status Date	• **Description:** The date of the latest available information on vital status, calculated from assessment dates, last info date, and date of death when available • **Implementation:** `add_time_to_last_vitalstatus()` • **Values:** Date format

7. Late Toxicity and Morbidity (only followup assessments)

Variable	Details
`max_value_*`	• Maximum grade of a specific toxicity or morbidity across all follow-up time points • **Calculation:** Maximum CTCAE grade value observed for the specific toxicity across all follow-up assessments • **Implementation:** `get_max_side_effects()` • **Values:** 0-5 (CTCAE grade scale) • **Example variables:** – `max_value_bladder_cystitis` – `max_value_vagina_stenosis` – `max_value_fistula` – `max_value_gastro_proctitis`
`max_timepoint_*`	• Follow-up timepoint at which the maximum grade of toxicity was observed • **Calculation:** Time point (e.g., "3m" for 3 months) at which maximum CTCAE grade was observed for each specific toxicity • **Implementation:** `get_max_side_effects()` • **Values:** Time point indicators (e.g., "3m", "6m", "12m", etc.) • **Example variables:** – `max_timepoint_bladder_cystitis` – `max_timepoint_vagina_stenosis` – `max_timepoint_fistula` – `max_timepoint_gastro_proctitis`
`overall_max_morbidity_ grade`	• Overall maximum morbidity grade across all toxicity endpoints and all time points • **Calculation:** The maximum value across all `max_value_*` variables for each patient • **Implementation:** `emii_add_max_morbidity()` • **Values:** 0-5 (CTCAE grade scale) • **Purpose:** Represents the highest toxicity experienced by a patient regardless of specific type or timepoint

The toxicity and morbidity data include various organ systems and toxicity types:

1. **Bladder toxicity**: Cystitis, frequency, incontinence, urgency, stenosis/stricture, bleeding
2. **Gastrointestinal toxicity**: Abdominal pain/cramping, constipation, diarrhea, incontinence, proctitis, bleeding (at different sites), stenosis/stricture
3. **Vaginal toxicity**: Bleeding, discharge, dryness, mucositis, stenosis
4. **Fistula**: Grading and localization (bladder, rectum, vagina, etc.)
5. **Lymphatic complications**: Edema (limb, trunk/genital), lymphocele, thromboembolic events
6. **Muscle/skeletal toxicity**: Fibrosis, fracture, back/pelvic pain
7. **Other toxicities**: Fatigue, hot flashes, insomnia, and other patient-specific issues