**Holistic Early Stage Prognostic Model: Data & Bulk**

Created by A Rodday, 5/8/23

Update on data decisions

* Waldeyer’s: There are only 3 who have Waldeyer’s only. They will be excluded.
* Inclusion/exclusion: age 18-65; cHL; early stage disease; no above & below diaphragm disease; no count of 0 EORTC nodal groups (looking above & below diaphragm), which also excludes Waldeyer’s ring only. MISSING >HALF OF CANDIATE VARIABLES. Registries further restricted to: treated with curative intent; not treated on a trial; diagnosed 1996 to 2019.
* Candidate variables:
  + Age, sex, stage (1v2), b symptoms, continuous measure of mediastinal bulk (using binary for now until we figure out units), histology (lymphocyte depleted, lymphocyte rich, mixed cellularity, nodular sclerosis, NOS), wbc count, lymphocyte count, hemoglobin, albumin, ESR, cervical nodal group, infraclavicular nodal group, mediastinal nodal group
* Outcome
  + Focus on PFS (275 events in trials; 82 events in registries)
  + Will not build OS model because of small number of events (150 in trials, 49 in registries)
    - WILL STILL REPORT TY OS
  + What to do with 8 that are missing outcomes from H10?
    - CENSOR AT 1 DAY

To discuss today

* Bulk
* Histology
* A-HIPI “True” external validation in Brazil Registry

Table 1. Bulk variables by study a

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **NCIC, n=284** | **H10, n=1803** | **H9U, n=728** | **RAPID, n=346** | **PMH, n=508** | **Iowa/ Mayo, n=244** | **Australia c, n=151** | **BC Cancer, 317** |
| **Mediastinal bulk** |  |  |  |  |  |  |  |  |
| Definition | Excluded if ≥ 1/3 max diameter | MT bulk ratio > .33 | MT ratio ≥ 0.35 | Excluded if ≥ 1/3 max. diameter | >1/3 max. diameter | >1/3 chest diameter | none | none |
| Variable type | binary (all no) | continuous (cm, ratio) | Categorical/ binary; continuous (cm, ratio) | continuous (cm) | binary | binary b | none | none |
| Mediastinal bulk, n (%) |  |  |  |  |  |  |  |  |
| Missing | 0 | 2 | 56 | 0 | 0 | 2 | 151 | 317 |
| No | 284 (100.0%) | 1254 (69.8%) | 375 (55.8%) | 378 (100.0%) | 239 (70.9%) | 214 (88.4%) | n/a | n/a |
| Yes | 0 (0.0%) | 543 (30.2%) | 297 (44.2%) | 0 (0.0%) | 98 (29.1%) | 28 (11.6%) | n/a | n/a |
| Mediastinal bulk in cm, mean (sd) | nmiss=284 | 5.7 (3.6), nmiss=435 | 6.9 (4.9), nmiss=391 | 1.7 (2.2), nmiss=31 | nmiss=508 | nmiss=244 | nmiss=151 | nmiss=317 |
| Mediastinal bulk as ratio, mean (sd) | nmiss=284 | 0.3 (0.1), nmiss=2 | 0.3 (0.2), nmiss=337 | nmiss=378 | nmiss=508 | nmiss=244 | nmiss=151 | nmiss=317 |
| **Non-mediastinal bulk** |  |  |  |  |  |  |  |  |
| Definition | Excluded if >10cm | ≥10 cm | none | ≥10 cm | > 10 cm | none | none | none |
| Variable type | binary (all no) | continuous (cm) | none | binary and continuous (cm) | binary | none | none | none |
| Non-mediastinal bulk, n (%) |  |  |  |  |  |  |  |  |
| Missing | 0 | 2 | 728 | 0 | 0 | 2 | 151 | 317 |
| No | 284 (100.0%) | 1782 (99.2%) | n/a | 375 (99.2%) | 328 (97.3%) | 228 (94.2%) | n/a | n/a |
| Yes | 0 (0.0%) | 15 (0.8%) | n/a | 3 (0.8%) | 9 (2.7%) | 14 (5.8%) | n/a | n/a |
| Non-mediastinal bulk in cm, mean (sd) | nmiss=284 | nmiss=25 | nmiss=728 | nmiss=31 | nmiss=508 | nmiss=244 | nmiss=151 | nmiss=317 |
| **Any bulk** |  |  |  |  |  |  |  |  |
| Definition | Excluded if ≥ 1/3 max diameter or >10cm | ? | MT ratio ≥ 0.35 or mass ≥10cm | ? | ? | >10 cm | > 5 cm | ≥10 cm |
| Variable type | binary (all no) | continuous (cm) | binary | continuous (cm) | continuous (cm) | binary | binary | continuous |
| Any bulk, n (%) |  |  |  |  |  |  |  |  |
| Missing | 0 | 4 | 39 | 378 | 171 | 1 | 8 | 1 |
| No | 284 (100.0%) | 1242 (69.1%) | 359 (52.1%) | n/a | 255 (75.7%) | 200 (82.6%) | 87 (60.8%) | 130 (41.1%) |
| Yes | 0 (0.0%) | 555 (30.9%) | 330 (47.9%) | n/a | 82 (24.3%) | 42 (17.4%) | 56 (39.2%) | 186 (58.9%) |
| Any bulk in cm, mean (sd) | nmiss=284 | 5.6 (3.2), nmiss=25 | nmiss=728 | 3.6 (1.9), nmiss=31 | 6.8 (3.6), nmiss=174 | nmiss=244 | nmiss=151 | nmiss=1 |

a Restricted to eligibility: age 18-65; cHL; early stage disease; no out of range lab values (albumin 1-6; ESR >=1; hemoglobin 5-16.5; lymphocyte count 0.1-5; WBC count 0.1-50); no disease above & below diaphragm; at least 1 nodal group per EORTC criteria (could be above or below diaphragm), which also excludes Waldeyer’s ring only. Registries further restricted to: treated with curative intent; not treated on a trial; diagnosed 1996 to 2019.

b Will get continuous data eventually

c For Australia, have a text field with location of bulk, but some missingness (10%) and inconsistently described. Also have max dimension of any bulk, but missing for anyone with bulk <5cm (so missing not at random) (73% missing).

**Histology**

How to classify those with “other” histology. For some we know it was cHL, for others, we don’t

These are the categories:

* 1="1LymphDepl" 2="2LymphRich" 3="3MixedCell" 4="4NodScler" 5="5NOS" 6="6cHL\_miss" 7="7Other\_cHL"   8="8NotCHL" 9="9NotHL" 10="10Unknown"
* For A-HIPI, we used multiple imputation for categories 6 & 7 (cHL missing histology and cHL other). Should we do that here as well?

NCIC

* 2 Interfollicular.
  + Annette, how should these be classified?
  + Annette says: I spoke with my Hematopathology colleagues about the 2 interfollicular cases from HD.6. It is hard to know what the pathologists decades ago were thinking. They lean towards calling NOS, although don't feel strongly.
  + SEEMS REASONABLE

 H10

* This study may not have been restricted to cHL, so a bit tricky to figure out what to do with the “others” and “unclassified”
* 11 initially categorized as “other” but they have no additional info. We may need to count them as “unknown” and exclude.
  + EXCLUDE THESE
* 41 initially categorized as “unclassified”. I had been counting these as NOS, but we could revisit if we aren’t confident they are cHL.
  + OK TO INCLUDE THESE AS CHL NOS. WE COULD FOLLOW-UP WITH STUDY TEAM ABOUT THIS (REIMAKERS(SP), MARC ANDRE, CC: JOHN RADFORD, ANDY, SKP)

RAPID

* 12 initially categorized as “other”. Don’t have any other info on them. But we do know this study was restricted to cHL.
  + SEND NICK THE PATIENT IDS FOR THESE. OTHERWISE, OK TO COUNT WITH OTHER AND IMPUTE
  + DID NOT DO CENTRAL REVIEW, BUT THE UK DOES REVIEW LOCALLY FOR HEMATOLOGY DIAGNOSES

**External Validation of A-HIPI in Brazil Registry**

1. Some patients with stage IIB disease from the Brazilian Registry were not included in the current analysis because they were not classified as advanced risk groups according to the GHSG risk group. Should these patients be included in the validation study as well?
   1. YES INCLUDE THEM.
2. When creating these new variables for hemoglobin, albumin, and lymphocyte count in the format for Holistic validation, those values that did not fulfill the permitted ranges were considered missing values. How should we handle missing data and values out of the permitted ranges?
   1. PRIMARY ANALYSIS EXCLUDED. SENSITIVITY ANALYSIS TO IMPUTE
3. We have 19 (2.6%) cases of early death (defined as deaths during ABVD treatment of patients who died of any cause during or up to 30 days after front-line treatment). What do you think about including them in a validation study?
   1. INCLUDE EARLY DEATH.