A Maximum Likelihood estimator of a Markov model for disease activity in chronic diseases that alternate between relapse and remission, for annually aggregated partial observations.

Sixten Borg. The Swedish Institute for Health Economics (IHE), Lund, Sweden.

Background

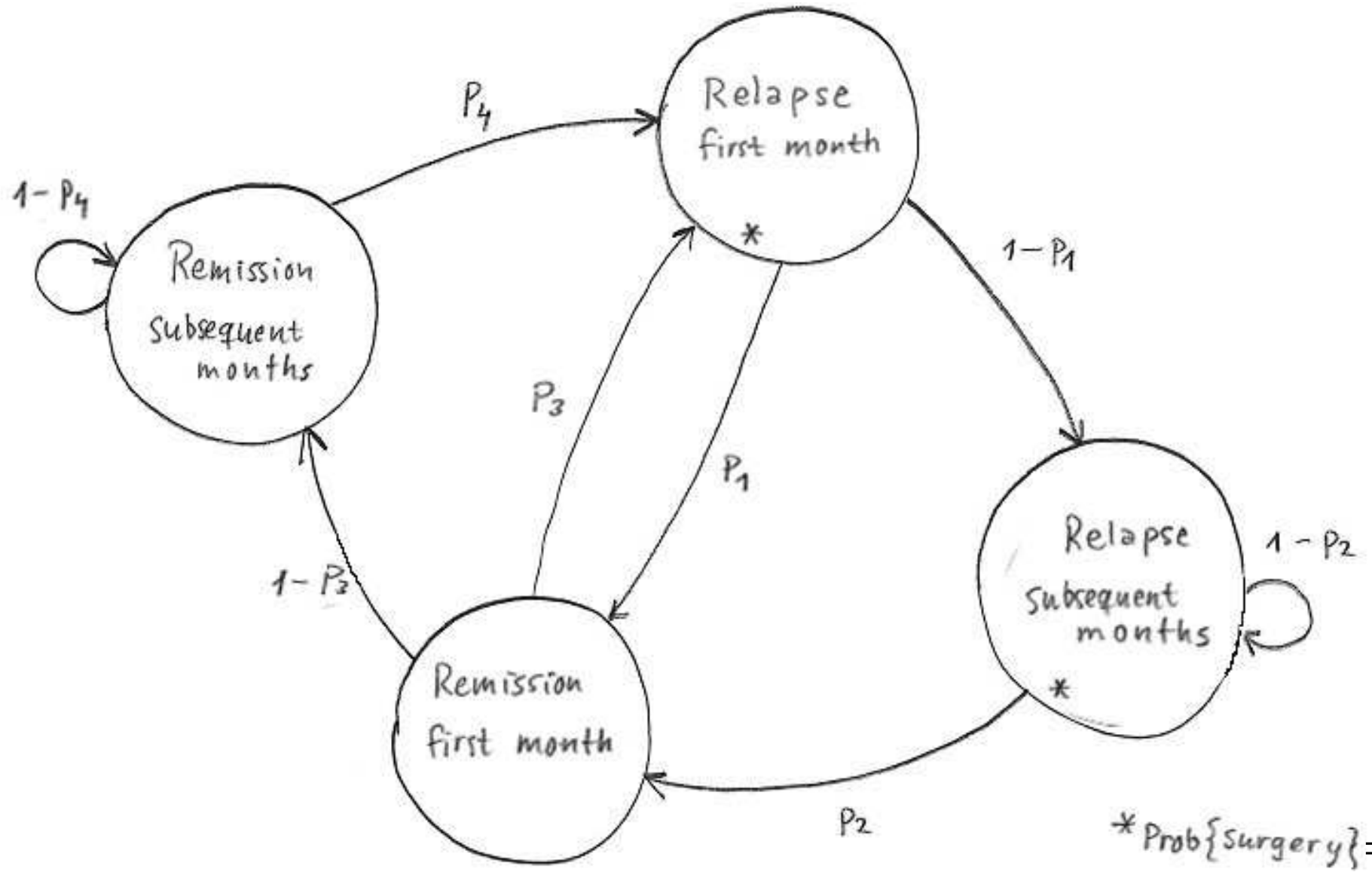
- Inflammatory Bowel Diseases, relapsing-remitting diseases.
- Crohn’s disease (CD)
  - Drug therapy, surgery. No cure.
- Ulcerative Colitis (UC).
  - Drug therapy, surgery.
  - ”Cure” through major surgery.

Purpose

- To obtain a model to study shortening of relapses or post-poning relapses.
  - Short cycle length (one month).
  - Time dependence.

Model

- Markov model S(t), with four states:
  - (1) first month of remission,
  - (2) subsequent months of remission,
  - (3) first month of relapse, and
  - (4) subsequent months of relapse.
- Surgery possible in states 3 and 4.


Relapse
first month
*
P_4
$1-P_4$
Remission
subsequent
months
$1-P_1$
P_3
P_1
Relapse
subsequent
months
*
$1-P_2$
$1-P_3$
Remission
first month
P_2
* Prob{surgery} = p_5

Data

- Yearly summaries of each patient's history:
  - No of relapses each year (V).
  - No of surgical operations each year (Z).
- No dates of surgery, or of relapse start/end.
- ~140 patients, ~10 years of follow-up.

Method

- Maximum-Likelihood estimator.
  - Function of transition probabilities and probability of surgery; (θ).
- Count over every possible pathway.
  - Some 17 million unique paths through model.

$$\Pr\{Z_l = z_l, V_l = v_l, S_{13} = s_{13} \mid S_1 = s_1, \theta\} =$$

$$= \sum_{s_2, s_3, \cdots, s_{12}=1}^{4} \left(\sum_{t_1, t_2, \cdots, t_{12}=0}^{1} \left(\begin{array}{l} I\{t_1 + t_2 + \cdots + t_{12} = z_l\} \cdot I\{v(s_1, s_2, \cdots, s_{12}) = v_l\} \cdot \\ \cdot \Pr\{S_2 = s_2, S_3 = s_3, \cdots, S_{13} = s_{13}, T_1 = t_1, T_2 = t_2, \cdots, T_{12} = t_{12} \mid S_1 = s_1, \theta\} \end{array} \right) \right)$$

Method

- Likelihood rewritten, to be more efficient:

$$\Pr\{Z_l = z_l, V_l = v_l, S_{13} = s_{13} \mid S_1 = s_1, \theta\} =$$

$$= (p_1)^{a_1}(1-p_1)^{b_1}(p_2)^{a_2}(1-p_2)^{b_2}\cdots(p_5)^{a_5}(1-p_5)^{b_5}$$

- There are some 12 thousand combinations of

$$s_1, s_{13}, a_1, b_1, a_2, b_2, \cdots, a_5, b_5, v, n$$

”profiles”, which can be determined in advance.

Method

- Likelihood determined through summation over all unique pathways / profiles.
- Numerical optimization of parameter values.
- Implemented in R with some components in C.

Results

- Behaviour of the estimator:
  - Estimator almost always converges.
  - Reasonable estimates with training datasets.
  - CD: estimates with good face validity.
  - UC: with curative surgery, *model inadequate*.

Results

- Optimization of the estimator:
  - Counting over pathways: estimator completes after about three hours.
  - After optimization: completes after about one minute.

Conclusions

- The estimator works.
- The existence of curative surgery calls for further development of the model and its estimator.

Possible future steps

- Turn into a proper R package?
- Generalize nature of aggregation and partial observation.
- Generalize underlying model.
  - Add state to deal with curative surgery?
  - Add states to improve ”time dependence”??
- Covariates, patient heterogeneity.