**Meeting Notes and Future Plans**

***Compiled by M.D. Ryser***

***December 17, 2016***

**1. Prerequisites** for successful estimation of natural history parameters and overdiagnosis in breast cancer (DCIS and invasive):

1. Need incidence with and without screening. Note: some information about no screening incidence can be extracted from screening trials and registries (given less than perfect screening sensitivity), but we may need to combine screening data with pre-screening incidence data.
2. We need to know the individual level screening patterns; BCSC will provide this level of information.
3. Need a natural history model that is identifiable.

**2.**  **Incidence of preclinical (latent) disease** poses modeling problems:

1. The constant hazard after a fixed initiation age does not seem meaningful. Results will be sensitive to initiation age.
2. One possibility is to use a Weibull to ensure that the hazard increases with age
3. Instead of a Weibull, we could also use a Gamma
4. Pinsky uses a cubic polynomial: 3 parameters and hence one step up from Weibull / Gamma
5. There is the possibility to use deconvolution (see earlier Pinsky paper) together with age-specific incidence of DCIS before screening was introduced…let’s discuss further
6. New ideas welcome…

**3. The Plan** (in 3 papers)

1. We start with the S-P/I-C model of Shen et al. and do the following: (i) fix the issue of incidence of preclinical disease (see also point 2 above); (ii) finish the structural identifiability analyses; (iii) move to individual-level likelihoods; (iv) simulate a synthetic dataset and investigate practical identifiability of the models; (v) fit it to the BCSC data.
2. Start working with the more detailed model (a la Pinsky), see Figure 1 below. (i) Build up from a model without indolent cancer and same sensitivity for DCIS and cancer, then do incremental steps by e.g. assuming different sensitivities for DCIS and invasive cancer, adding indolent invasive cancer, and both; (ii) Do as much formal structural identifiability work as possible; (iii) simulate synthetic data and investigate practical identifiability and estimability of the models; (v) fit to the BCSC data with individual-level likelihoods.
3. Extend the model to look at more detailed natural history pathways, accounting for (i) ER status; (ii) nuclear grade; (iii) low vs high risk as in active surveillance trials. Repeat the steps from paper 2 and fit to BCSC data.

**4. Data preparation**. BCSC data will be curated in consideration of the following aspects:

* Focus on women for whom we have information on their first screen
* Possibly further sub-setting to make sure we have only small gaps in the patient-level history (can be tricky because they may get additional screens outside the registry etc)
* ER status
* Grade
* Size
* Screening modality
* More?

**5. Analyses**. We need to try and extend the differential algebra methods for identifiability to non-Markovian processes, in particular for the incidence of preclinical disease. I will talk to Marisa, Andrew and Rafael about this (after thinking about it myself).

**6. Other points for future discussions**

* Do we need to condition on being asymptomatic on the baseline screen when deriving the likelihood for screening data?
* Briefly touched upon the general problem that people in surveillance cannot die form cancer – I would like to get back to this issue and how to deal with it.
* Sensitivity is not constant over time. It depends on tumor features such as size. How can we accurately capture this? At the same time, clinical incidence increases with size, too. I am wondering if we may want to consider a birth-death model for growth of DCIS with a one-hit model for progression to invasive disease. This could give us an underlying size distribution that could inform both screening sensitivity and the clinical detection hazard.
* We may want to use autopsy data (age-stratified) in the way Etzioni et al did in AJE 1998 for prostate cancer. At the very least, autopsy data can constrain some of the inferences
* It will be good to collaborate on the risk communication part as this is a part of both the overdiagnosis and the K99/R00 grants.

Macintosh HD:Users:mdryser:ownCloud:DCIS:Etzioni:Figure_1.pdf

**Figure 1: Basic model of DCIS and invasive cancer cancer.**