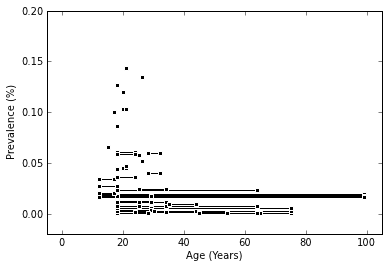
**Cannabis Dependence**

The meta-analysis of data from a systematic review of the prevalence of cannabis dependence has a clear age-specific prevalence, providing an example of the importance of spline models.

Symptoms associated with cannabis dependence are compulsive use and difficulty with abstinence [1]. The American Psychiatric Association recognizes cannabis dependence as fulfilling three of the following seven criteria:

* tolerance
* withdrawal
* substance is taken in larger amounts or over longer period than intended
* persistent desire or unsuccessful efforts to control substance use
* great deal of time is spent to obtain use or recover from effects of substance
* important social, occupational or recreational activities are reduced because of substance use
* continued substance use despite knowledge of physiological or psychological problems induced by substance use [2]

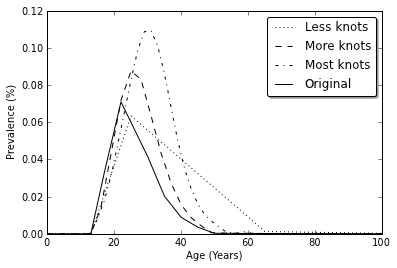
Fifteen studies were identified for cannabis dependence prevalence, covering 7 countries. Since there are few data available on cannabis dependence, data points for the prevalence of cannabis use are also included in the analysis, increasing the prevalence data to 101 studies covering 92 countries. To avoid overestimation with the inclusion of cannabis use, a study-level covariate distinguishes between cannabis use and dependence data. Study level covariates will be discussed in more detail in Chapter **??**. However, for simplicity the following graphs show cannabis dependence prevalence data only.



Above: Global data for cannabis dependence.

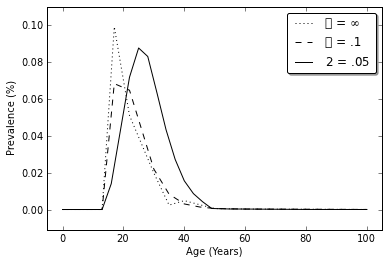
As discussed in Chapter **4**, age-specific hazards are modeled with spline models. A spline model is any piecewise polynomial function. Knots partition the age range into intervals. With ample data and clear age patterns, models will not be very sensitive to choice of knots. However, when working with sparse and noisy data, the number and location of knots are important decisions as they can influence the model results substantially. When this is the case, the number of knots and locations should be chosen a priori using expert knowledge concerning the disease being modeled.

The original model for cannabis dependence has knots at the ages of 0, 13, 17, 22, 30, 35, 40, 45, 50, 60, and 100. Three other models are shown. The ‘less knot’ model limits knots to ages 0, 13, 25, 65, and 100. The ‘more knot’ model has knots every three years (0, 13, 16, 19, …, 43, 46, 49) between ages 13 and 50 with additional knots at 55, 60, 80 and 100. The ‘most knot’ model has knots every two years during the ages of 13-66 (knots at ages 0, 13, 15, 17, …, 61, 63, 65, and 100). As seen from the figure below, choosing too few or too many knots can influence the model substantially.



Above: Prevalence of cannabis dependence in original model and models with ‘less, more and most knots’ .

A penalized spline model with a smoothing parameter is another solution to knot selection. The model includes more knots but adds a penalty to discourage using more knots than necessary for the data. The smoothing parameter controls the roughness of the estimating function and the fidelity of the data, making knot selection matter less as shown in the following figure.



Above: Original model and a penalized splines model ith a smoothing parameter σ.

[1] Coffey, C, Carlin, JB et. al. 2002. “Cannabis dependence in young adults: an Australian population study”. *Addiction*. 97(2)187-194.

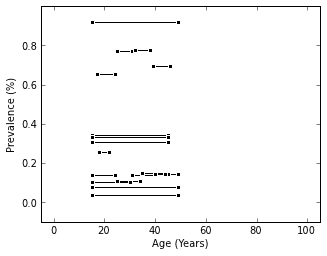
[2] American Psychiatric Association. 1994. “Criteria for Substance Dependence”. *Diagnostic and Statistical Manual of Mental Disorders, 4th Ed*. Washington DC: American Psychiatric Association. p. 197

**Unclear age pattern, requiring expert priors: Premenstrual Syndrome**

Epidemiological data without clear age-specific patterns are a recurring theme in the GBD 2010 study. Unclear age patterns make expert priors essential. However, such cases are very sensitive to the choice of prior assumptions, as shown in the following example of premenstrual syndrome (PMS) in Western Europe.

PMS is a common cyclic disorder that affects women of reproductive years during the period between ovulation and the onset of menses. More than 200 behavioral, psychologic and physical symptoms have been associated with PMS, the most common being irritability, tension, depression, bloating, weight gain and food cravings [1, 2]. There is no known cause or consistent treatment [3].

A meta-analysis of data from a systematic review on the descriptive epidemiology of premenstrual syndrome yielded 74 prevalence data points, of which 18 were from Western Europe. As seen from the figure below, the data are noisy, with overlapping and heterogeneous age groups that show no clear age pattern.

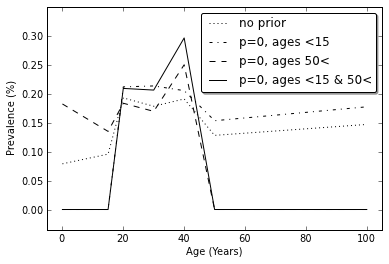


Above: Prevalence data for women with premenstrual syndrome in Western Europe.

**Priors on level and monotonicity**

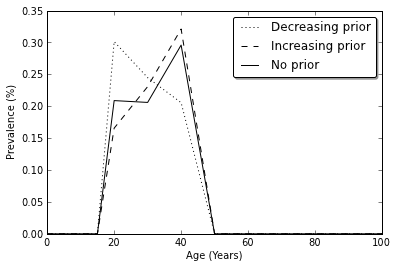
With no clear guidance from the data, informative priors to identify an age pattern are a critical part of the modeling process although they may have unintended effects, as discussed in Chapter **5**. To illustrate the effects of priors, an age standardized negative-binomial mixed effects spline model is used without invoking the compartmental model from chapter ??.

Looking at the data in the figure above, no data exists before age 15 or after age 50. Since PMS is a disorder related to the cycles of the female reproductive system, it is obvious that the data outside this age range are not present for biological reasons. However, this information is unknown to the model and without a prior limiting the age range to 15-60, the spline model estimates prevalence for the entire age range. Here expert knowledge needs to infor the model that prevalence should not be expected outside of the ages 15-50.



Above: Prevalence of premenstrual syndrome without (a) and with (b) a prior on age range

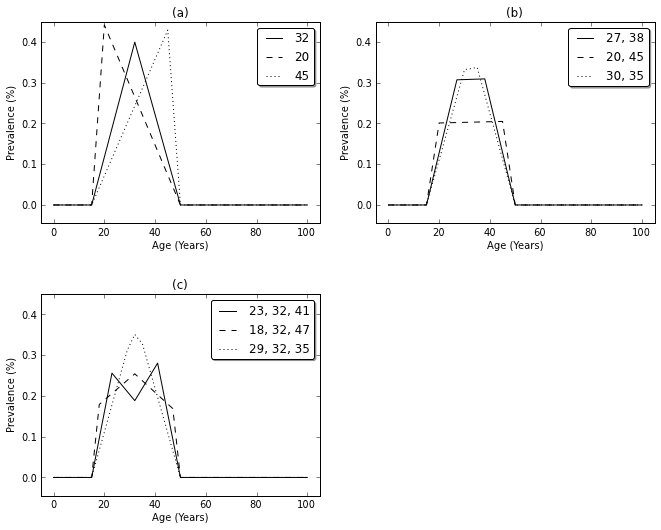
Another common prior for age patterns is the belief that the function is increasing or decreasing over a certain age range. As seen in the figure below, priors on monotonicity between the critical ages of 25 and 40 have a big effect on the prevalence estimate for Western Europe.



Above: Between the ages of 25-40, the prior on monotonicity makes a large impact on the prevalence estimates for women in Western Europe with premenstrual syndrome.

**Knot location**

As previously discussed in Chapter **4** and **(cannabis dependence example)**, age-specific hazards are modeled with spline models, using knots to partition the age range into intervals. With ample data and clear age patterns, models will not be very sensitive to choice of knots. However, with sparse and noisy data without a clear age pattern, the number and location of knots can influence the model results substantially. Choosing the number and location of knots a priori using expert knowledge allows the user to determine critical features of the model.



Above: All panels have knots at 0, 15, 50, 100 and vary the number and location of knots between the ages of 15 and 50 to show the sensitivity of knot selection sparse and noisy data without a clear age pattern. Even with 1 knot, the placement at age 20, 32 or 45 gives markedly different results. Figure (b) uses 2 knots and varies their locations at ages 27 and 38, 20 and 45, 30 and 35 while Figure (c) uses 3 knots at locations [23, 32, 41], [18, 32, 47] and [29, 32, 35].

[1] Dickerson LM, et. al. 2003. “Premenstrual Syndrome”. *American Family Physician*. 67(8):1743-1752.

[2] Singh BB, et. al. 1998. “Incidence of premenstrual syndrome and remedy usage: a national probability sample study”. *Altern Ther Health Med*. 4(3):75-79.

[3] Goodale IL, et. al. 1990. “Alleviation of Premenstrual Syndrome Symptoms with the Relaxation Response”. *Research and Resources*. 75(4):649-655.

[4] <<http://winthrop.ihme.washington.edu/dismod/summary/32404>>. Accessed 17 Apr 2012.

**Dealing with geographical variation: Hepatitis C**

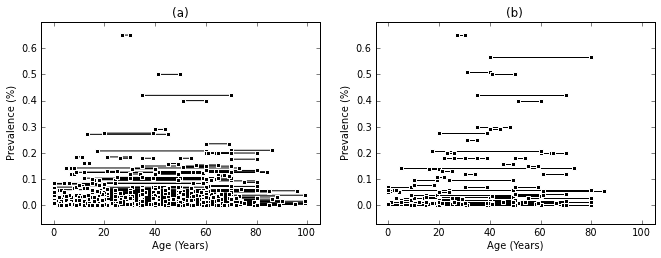
Hepatitis C is a viral infection that attacks the liver. In a small portion of acute cases, the body can eliminate the virus; however the majority of acute cases develop into chronic infections [1]. Chronic infections cause liver damage and may develop to end stage liver disease or cirrhosis. Few, if any, chronic cases experience symptoms and only one third of acute cases are symptomatic and jaundice. Chronic symptoms are nonspecific, intermittent and mild with the most common symptom being fatigue [2, 1]. Common symptoms for severe and advanced disease stages include nausea, dark urine and jaundice [1]. Since hepatitis C infections are asymptomatic, diagnosis requires laboratory testing for both hepatitis antibodies (anti-HCV) and the hepatitis virus (HCV RNA) [2]. There is no vaccination for hepatitis C, but therapy can prevent advanced liver disease [3].

Compared to other countries in the region, Egypt has a high hepatitis C prevalence. In an attempt to treat endemic schistosomiasis, a common parasitic worm that affects the urinary tract, gut and liver, the Egyptian Ministry of Health launched widespread injection-based treatment throughout 1950-1980. While there were improvements in schistosomiasis-induced mortality, recycled needles and poor needle sterilization infected many with hepatitis C [4, 5, 6]. The spatial variations of hepatitis C in North Africa and the Middle East make it an excellent example of hierarchical random effects modeling.

Random effects modeling detects systematic differences among different hierarchies, or levels, of data. The spatial hierarchy in the GBD 2010 study uses countries nested in regions nested in super-regions. There are 21 regions defined by demographic and epidemiological similarities that are further clustered by 7 super-regions.

The analysis of hepatitis C uses data on the prevalence of person who have Hepatitis C antibodies to compare hepatitis C infection levels globally. Incomplete data or data from high-risk populations such as health workers were excluded [3].

Hepatitis C prevalence in Egypt is 40 times that of Jordan even though they are part of the same GBD world region North Africa and the Middle East (Figure 13.1),



Above: Prevalence data from systematic review of hepatitis C in Jordan (panel a) and Egypt (panel b).

The analysis uses an age-standardizing hierarchical random effects generalized negative binomial spline model to estimate prevalence. The hierarchical random effects allow the model to capture variation within the region of North Africa and the Middle East. Looking at the table below, Egypt (EGY) has significantly higher prevalence than the other countries in the region. Figure **?.?**, confirms this as the estimate for Egypt is much above the regional average.

Hepatitis C prevalence estimations from random effects model}

Country & Posterior Mean & Lower 95\% HPD & Upper 95\% HPD

EGY & 1.86 & 1.6, & 2.2 \\

JOR & -0.6 & -1.0, & -0.1 \\

SAU & -0.77 & -1.2, & -0.4 \\

IRQ & 0.05 & -0.4, & 0.5 \\

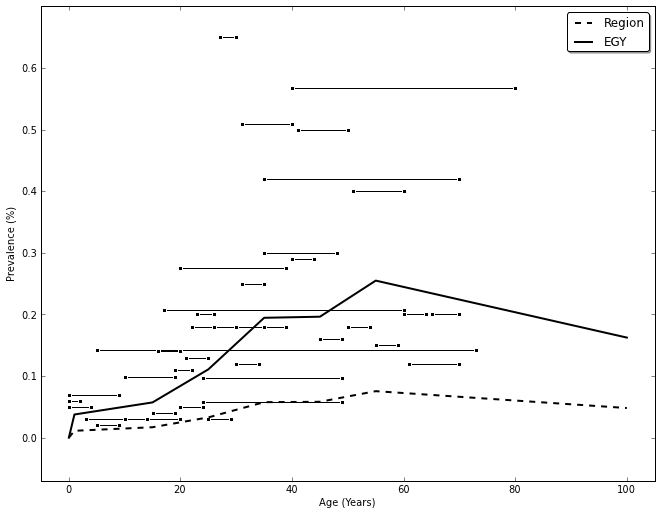
IRN & 0 & -0.4, & 0.5 \\

YEM & 0.04 & -0.3, & 0.5 \\

TUR & -0.32 & -0.6, & 0.1 \\

SYR & -0.16 & -0.6, & 0.3 \\

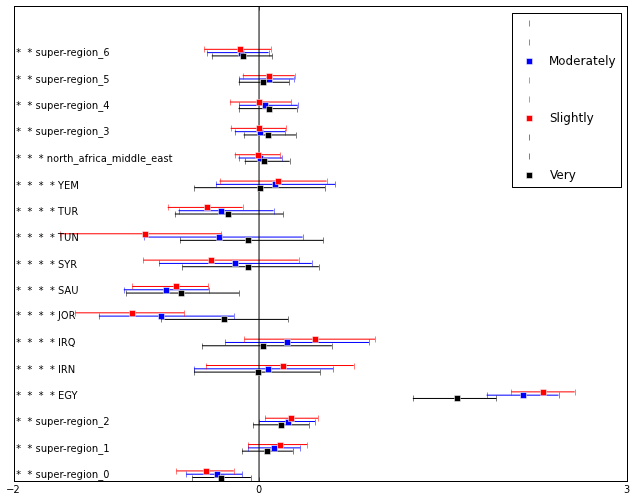
TUN & -0.18 & -0.6, & 0.4 \\



Above: The 1990 estimate of hepatitis C prevalence for men the region of North Africa and Middle East and the country of Egypt. These estimates only use 2 levels in the hierarchal random effects model—region and country.

In such noisy data, placing a prior on the dispersion of the data informs the model of the data heterogeneity. This allows the model to infer how dispersed the random effects are between geographic regions, and hence quantify the uncertainty in the geographic regions for which no data is available. Changing the prior for the heterogeneity of the global data has effects on the country level as seen in the figure below. Random effects modeling detects within sample variation and true variation that cannot be explained by a covariate. Therefore, a change in the prior on global heterogeneity changes the level of variation and thus the random effect size. As seen in the figure below, when the prior on global heterogeneity is ‘very’, the estimates are compressed.

better discussion of compression needed.



Above: The 1990 estimate of hepatitis C prevalence for men with different priors on global heterogeneity, (a) ‘slightly’, (b) ‘moderately’ and (c) ‘very’). Four levels (global, super-region, region, country) were used in the hierarchal model random effects model. Notice that changes for in the prior on global data heterogeneity affects the estimates at the country level (\*\*\*\*) the most.

[1] Hoofnagle, JH. 1997. ”Hepatitis C: The Clinical Spectrum of Disease”. *Hepatology*. 15S-20S

[2] Ghany, MG, et al. 2009. “Diagnosis, Management, and Treatment of Hepatitis C: An Update”. *Hepatology*. 49(4):1335-1374.

[3] Hanafiah, KM, et. al. 2012. “Global epidemiology of hepatitis C virus infection: New estimates of age-specific antibody to hepatitis C virus seroprevalence”. *unpulished*.

[4] Frank, C, et al. 2000. “The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt”. *The Lancet*. 355:887-891.

[5] Mezban, ZD and Wakil AE. 2006. “Hepatitis C in Egypt”. *American Journal of Gastroenterology*.

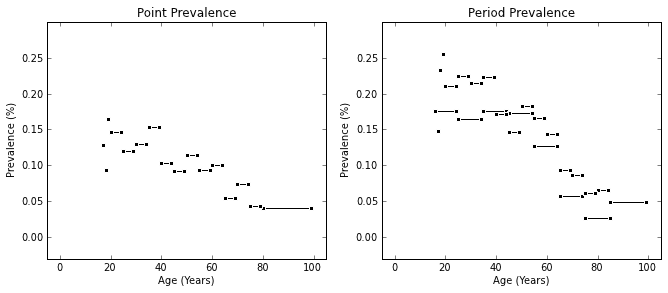
[6] Strickland, GT. 2006. “Liver disease in Egypt: Hepatitis C superseded schistosomiasis as a result of iatrogenic and biological factors”. *Hepatology*. 43(5):915-922.

**Study quality fixed effects: Anxiety**

Fixed effects explain the bias and variation of noisy measurements in terms of demographic, epidemiologic and study-specific variables. Unlike random effects that vary by geographic unit, fixed effects have covariates that differ by study or by country and year. Frequently in meta-analysis of mental disorders, such as anxiety, studies use different recall periods in the measurement of epidemiologic rates. In this case, fixed effect study-level covariates explain the bias of the measurements resulting from different recall periods.

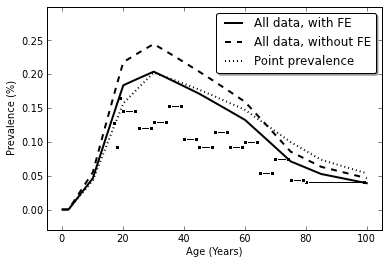
Anxiety disorders include at least eight separate conditions each characterized by prominent anxiety at a level which interferes with daily life. As there is a lot of co-morbidity between individual anxiety disorders, the expert group decided to model all anxiety disorders together as a single condition for GBD. Not all anxiety disorders manifest in similar ways. While generalized anxiety disorder is typically marked by persistent worry, panic disorder is usually characterized by intense fear for discrete periods of time [1].

Anxiety disorders do not have a consistent recall period for the measurement of epidemiological rates. Therefore the data from meta-analysis has studies with measurements of point prevalence and period prevalence (i.e. 6-month or past year prevalence). Due to the high remission rate for anxiety disorders, period prevalence is typically higher than point prevalence as seen in the figure below.



Above: Data on female anxiety disorders collected in systematic review for 2000-2010 in Australasia. Some studies measured point prevalence or period. Period prevalence is typically higher than point prevalence because of the high remission rate for anxiety disorders.

Excluding period prevalence measurements reduces the quantity of data and produces results that are generally much lower than estimates using all of the data without fixed effects. Using a fixed effect study-level covariate to explain the systematic bias and variation resulting from different recall periods lowers the prevalence estimate as seen below.



Above: Prevalence estimates for anxiety disorders in 2005 for women in Australasia, with point prevalence data shown. Notice that estimates based on all data without fixed effects are higher than point prevalence estimates.

[1] American Psychiatric Association. 2000. “Criteria for Substance Dependence”. *Diagnostic and Statistical Manual of Mental Disorders, 4th Ed Text Revision*. Washington DC: American Psychiatric Association. p. 429-484.

**Chronic Kidney Disease (CKD)**

Chronic Kidney Disease, commonly known as CKD, is the slow and progressive loss of kidney function [1]. There are five stages to the disease, with the final stage being kidney failure (end-stage renal disease--ESRD). The most common causes of CKD are diabetes and high blood pressure [1]. Damage to kidneys is usually permanent, but treatment and life style changes can slow the disease progression [2].

Since symptoms develop late in the disease, the progress of CKD is determined by the level of Glomerular filtration rate (GFR), an estimate of the flow of filtered fluid through the kidney [2, 3]. Patients with Stage 1 CKD have GFR levels above 90 mL/min/1.73 m$^2$. At this stage, kidney damage has already begun even though GFR remains at normal levels. Stage 2 CKD is defined as GFR between 60 and 90 mL/min/1.73 m$^2$. At Stage 3 CKD, a few symptoms appear with laboratory abnormalities in other organ systems as the GFR decreases to levels between 30 and 60 mL/min/1.73 m$^2$. When the GFR decreases to levels between 15 and 30 mL/min/1.73 m$^2$, Stage 4 CKD patients have mild symptoms and many organ systems display lab abnormalities. Stage 5 CKD is kidney failure, defined as having GFR levels below 15 mL/min/1.73 m$^2$ and not on being on renal replacement therapy, such as dialysis [3]. At this stage, the kidneys no longer function and the patient needs dialysis or a kidney transplant to live [1].

This example focuses on ESRD dialysis patients because it has four different types of data (prevalence, incidence, remission and with-condition mortality). “For analysis, 5315 data points from 129 studies/registry reports were included, as well as 349 data points from the Genitourinary Expert Group for country-years with a known lack of dialysis or transplantation with no other data available, for a total of 5664 data points representing 161 countries in all 21 regions. Transplantation incidence among the prevalent dialysis population was used as a proxy for remission.” [4] The model combines hemodialysis and peritoneal dialysis for analysis. Data without measures of uncertainty (effective sample size, standard error or 95% confidence intervals) and duplicated data from different sources are excluded from analysis.

\begin{table}[h]

\begin{center}

\caption{ Frequency of dialysis morbidity data types by region used in estimation }

\label{tab:CKD\_data}

\begin{tabular}{|l|c|}

\hline

Region & Prevalence & Incidence & Remission & Mortality & Total \\

\hline

North America, High Income & 277 & 554 & 241 & 240 & 1312 \\

Europe, Western & 362 & 694 & 65 & 10 & 1131 \\

Australasia & 332 & 403 & 144 & 219 & 1098 \\

Asia Pacific, High Income & 161 & 168 & 16 & 26 & 371 \\

Asia, Southeast & 138 & 131 & 14 & 0 & 283 \\

Latin America, Southern & 112 & 113 & 20 & 0 & 245 \\

Asia, East & 107 & 105 & 7 & 0 & 219 \\

Europe, Central & 102 & 86 & 18 & 10 & 216 \\

North Africa/Middle East & 76 & 55 & 17 & 2 & 150 \\

Europe, Eastern & 53 & 53 & 13 & 0 & 119 \\

Sub-Saharan Africa, West & 21 & 18 & 51 & 0 & 90 \\

Asia, South & 35 & 29 & 8 & 0 & 72 \\

Latin America, Tropical & 42 & 10 & 6 & 6 & 64 \\

Sub-Saharan Africa, East & 12 & 9 & 39 & 0 & 60 \\

Latin America, Central & 24 & 21 & 12 & 1 & 58 \\

Caribbean & 15 & 9 & 33 & 0 & 57 \\

Oceania & 13 & 13 & 18 & 0 & 44 \\

Sub-Saharan Africa, Central & 9 & 9 & 18 & 0 & 36 \\

Sub-Saharan Africa, Southern & 6 & 2 & 15 & 0 & 23 \\

Asia, Central & 2 & 1 & 6 & 0 & 9 \\

Latin America, Andean & 5 & 2 & 0 & 0 & 7 \\

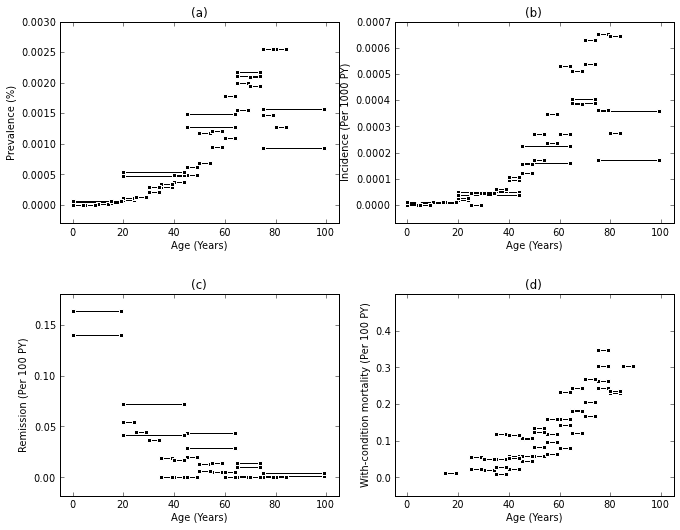
Total & 1904 & 2485 & 761 & 514 & 5664 \\

\hline

\end{tabular}

\end{center}

\end{table}



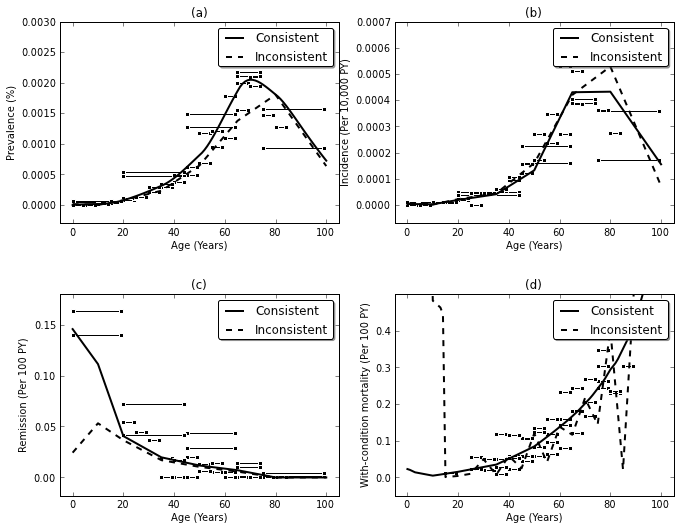
Above: Data for 2005 Australasia ESRD dialysis has four data types—prevalence, incidence, remission and with-condition mortality. The model combines hemodialysis and peritoneal dialysis for analysis. Data without measures of uncertainty (effective sample size, standard error or 95% confidence intervals) and duplicated data from different sources are excluded from analysis.

**Consistent v. Inconsistent Results**

As discussed in Chapter **1.5**, epidemiologic parameters, such as incidence, prevalence, remission, and with-condition mortality, are related by a logical requirement of internal consistency. A prevalent case can only exist if there was a past incidence event and the current number of prevalence cases can be determined from past prevalent cases, new incident cases, deaths and remissions. Modeling the parameters simultaneously produces a best estimate and plausible uncertainty bounds for incidence and prevalence that are internally consistent estimates for a single time, place and sex.

The figure below compares the consistent and inconsistent results for Australasian males with dialysis treatment for CKD in 2005. The consistent model estimates prevalence, incidence, remission and with-condition mortality simultaneously. The inconsistent model models each parameter individually.

Perhaps more of a discussion of the differences??



Above: The figure compares the consistent and inconsistent results for Australasian males with dialysis treatment for CKD in 2005. The consistent model estimates prevalence, incidence, remission and with-condition mortality simultaneously. The inconsistent model models each parameter individually.

[1] National Kidney Foundation. 2012. “About Chronic Kidney Disease”. <http://www.kidney.org/kidneydisease/aboutckd.cfm> Accessed Apr 3 2012.

[2] American Kidney Fund. 2010. “Chronic Kidney Disease (CKD)”. <http://www.kidneyfund.org/kidney-health/kidney-problems/chronic-kidney-disease.html> Accessed Apr 4 2012

[3] National Kidney Foundation. 2002. “KDOQI Clinical Practice Guidelines for Chronic Kidney Disease: Evaluation, Clasification, and Stratification. Part 4. Definition and Classification of Stages of Chronic Kidney Disease”. <http://www.kidney.org/professionals/kdoqi/guidelines_ckd/p4_class_g2.htm> Accessed Apr 3 2012.

[4] Wulf, Murray, and Naghavi. 2012. “2010 GBD Estimation Strategy Report for Chronic Kidney Disease”.

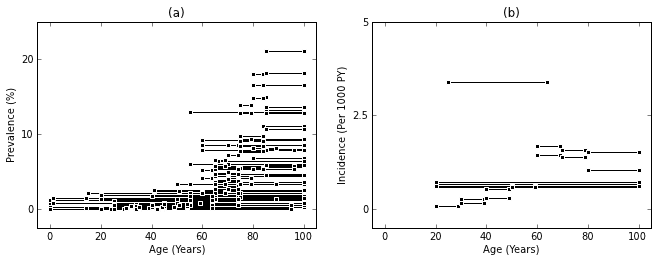
**Overlapping, heterogeneous age groups: Atrial Fibrillation**

As discussed in Chapter **7**, atrial fibrillation research has no standard set of age groups for study. Therefore the meta-analysis of data from systematic review on the descriptive epidemiology parameters of atrial fibrillation provides an excellent example of overlapping and heterogeneous age groups.

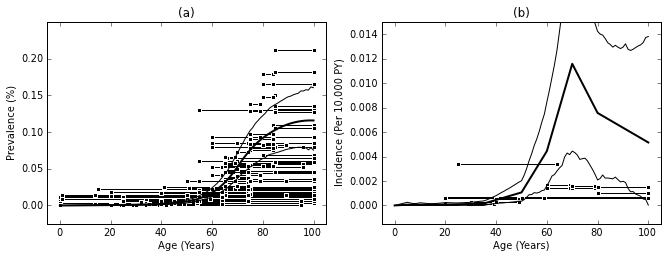
The most common type of cardiac arrhythmia is atrial fibrillation [1, 2]. Chaotic and irregular heart rhythms originating in the atria cause poor blood flow to the body. Atrial fibrillation episodes may be occasional, only lasting a few minutes or hours, or chronic if the heart rhythm is always abnormal [3]. Symptoms include heart palpitations, lack of energy, dizziness, shortness of breath and chest discomfort, although some cases of atrial fibrillation are symptomless [2]. Atrial fibrillation may occur at any age with increasing risk at older ages [4, 1]. It is uncommon in children [4].

For analysis, atrial fibrillation in Western Europe has 200 prevalence data points and 29 incidence data points [5].

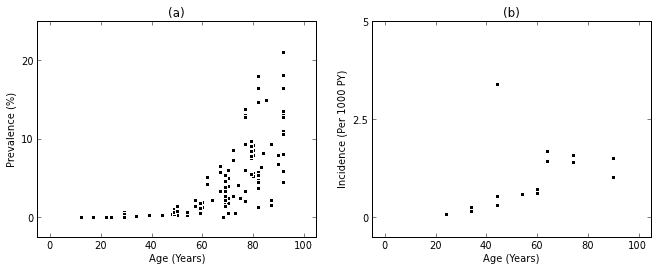
As seen from the data below, atrial fibrillation has heterogeneous and overlapping age groups. Without access to the microdata needed to recreate homogeneous age groups, an alternate approach must be used. As a solution to the heterogeneous age groups, *age-standardizing* adds age-weights to the age-specific rate according to population structure. The age-standardizing model uses a common age pattern for all studies so that the age-weights are the same for all age groups. Please refer to Chapter **7.5-6** for more details.



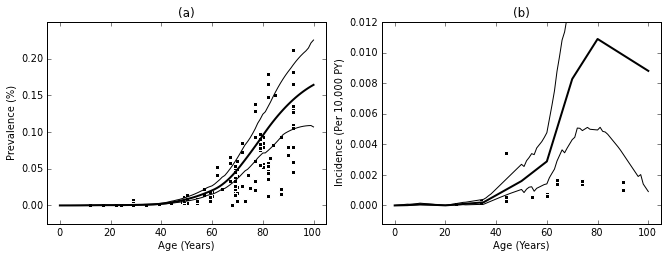
Above: Data for Western Europe males with atrial fibrillation is an excellent example of heterogeneous and overlapping age groups.



Above: Estimates of prevalence and incidence of atrial fibrillation in males in Western Europe in 1990. As discussed in Chapter **7.2**, the simplest approach to modeling heterogeneous age groups is to apply each age-specific rate measurement to the midpoint of the age interval.

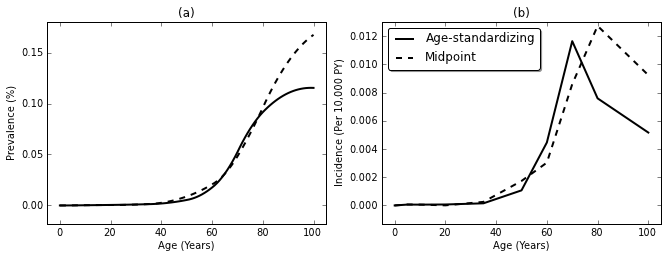


Above: To eliminate heterogeneous and overlapping age groups, only the midpoints of age groups are used in the midpoint method. Here the midpoints of the age groups for Western Europe males with atrial fibrillation data are shown.



Above: Using the midpoint method, data were fit by applying the age-specific rate to the midpoint of the age group for estimated prevalence and incidence for Western Europe males with atrial fibrillation in 1990

Comparing the two methods--discussion here about the differences?



Above: Comparison of the age-standardizing and midpoint methods for estimated prevalence and incidence for Western Europe males with atrial fibrillation in 1990.

[1] Rich, MW. 2009. “Epidemiology of atrial fibrillation”. *J Interv Card Electrophysiol*. 25:3–8.

[2] Rho, RW and Page RL. 2005. “Asymptomatic Atrial Fibrillation”. *Progress in Cardiovascular Diseases*. 48(2)79-87

## [3] American College of Cardiology Foundation. 2006. “ACC/AHA/ESC 2006 Guidelines for the Management of Patients With Atrial Fibrillation—Executive Summary”. *J Am Coll Cardiol*. 48:854-906

[4] Radford, DJ and Izukawa T. 1977. “Atiral Fibrillation in Children”. *Pediatrics*. 59(2)

[5] DisMod3. “Summary of Model, Amount of Data by Region and Type”. <<http://winthrop.ihme.washington.edu/dismod/summary/32458>

**Bipolar disorder**

The meta-analysis of data from a systematic review on the descriptive epidemiology of bipolar disorder provides an excellent example of the effects of informative priors on levels of age-specific incidence and remission hazards.

Bipolar disorder is a mental disorder that causes mood swings fluctuating between euphoric highs called manic episodes and depressive lows, interspersed by periods of residual symptoms {American Psychiatric Association, 2000 #31}. Manic episodes may last from days to months, causing personal, social and work-related problems [2, 3]. Mood swings may occur as infrequently as yearly or as frequently as several times a day [4]. Extreme behavior changes accompany mood changes, and it is not uncommon for sleeping, eating or activity patterns to change with manic and depression episodes [3]. There is no clear cause for episodes, but life changes, medications and sleeplessness may trigger manic periods [2]. While there is no cure, treatment helps manage mood swings and related symptoms [4].

The disease modeling of bipolar disorder was based on literature describing it as a chronic illness with little or no complete remission {American Psychiatric Association, 2000 #31}. This approach differed to the modeling of chronic episodic mood disorders like major depressive disorder where an estimate of disease duration rather than remission is required.

Thirty-two studies were identified for prevalence, covering 22 countries in 11 GBD world regions (Table 1). Two studies were identified for incidence, both from the United States of America. Seven studies were identified for excess mortality, from 5 high income countries. We found no studies reporting on complete remission (equivalent to cure rather than a temporary reduction in symptom levels as clinicians tend to define ‘remission’) from bipolar disorder as defined by GBD, consistent with the description in the literature that there is no cure for bipolar disorder.

\begin{table}[h]

\begin{center}

\caption{ Frequency of bipolar data types by region used in estimating disease parameters}

\label{tab:CKD\_data}

\begin{tabular}{|l|c|}

\hline

Region & Prevalence & Incidence & Remission & Mortality \\

\hline

Europe, Western & 24 & 0 & 0 & 13 \\

North America, High Income & 19 & 2 & 0 & 4 \\

Australasia & 16 & 0 & 0 & 0 \\

Asia, East & 12 & 0 & 0 & 0 \\

Asia Pacific, High Income & 8 & 0 & 0 & 2 \\

Latin America, Tropical & 4 & 0 & 0 & 0 \\

Sub-Saharan Africa, East & 2 & 0 & 0 & 0 \\

North Africa/Middle East & 2 & 0 & 0 & 0 \\

Latin America, Southern & 2 & 0 & 0 & 0 \\

Latin America, Central & 2 & 0 & 0 & 0 \\

Europe, Eastern & 2 & 0 & 0 & 0 \\

Sub-Saharan Africa, West & 0 & 0 & 0 & 0 \\

Sub-Saharan Africa, Southern & 0 & 0 & 0 & 0 \\

Sub-Saharan Africa, Central & 0 & 0 & 0 & 0 \\

Oceania & 0 & 0 & 0 & 0 \\

Latin America, Andean & 0 & 0 & 0 & 0 \\

Europe, Central & 0 & 0 & 0 & 0 \\

Caribbean & 0 & 0 & 0 & 0 \\

Asia, Southeast & 0 & 0 & 0 & 0 \\

Asia, South & 0 & 0 & 0 & 0 \\

Asia, Central & 0 & 0 & 0 & 0 \\

Total & 93 & 2 & 0 & 19 \\

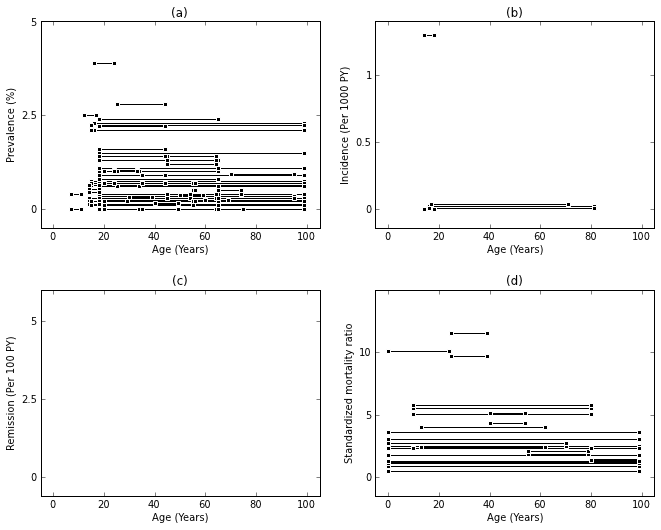
\hline

\end{tabular}

\end{center}

\end{table}

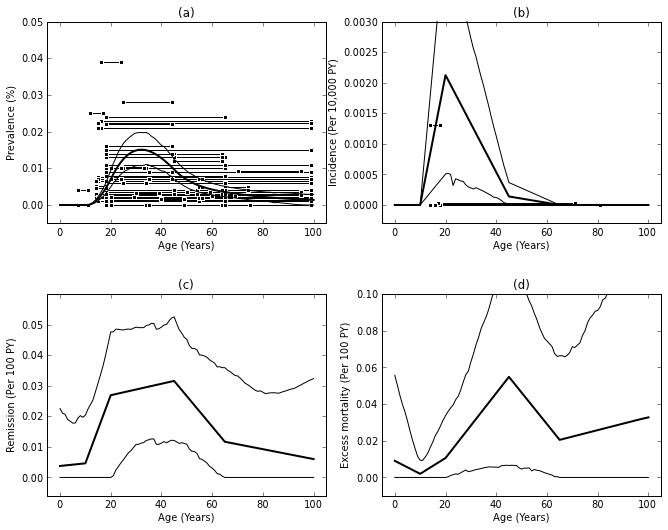
There was considerable variability in the data. For instance, estimates were reported across multiple age groups (specific or broad), they were based on different coverage areas (community or national) and they were either sex specific or for both sex combined. This data has been summarized in greater detail elsewhere {Ferrari, 2011 #1574; Ferrari, 2011 #1581}.



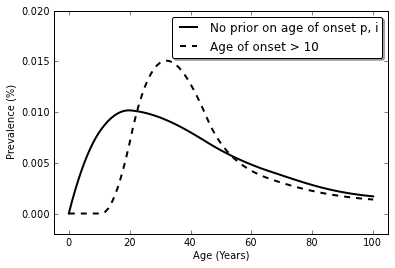
Above: Data included for the modeling of bipolar disorder. Thirty-two studies were identified for prevalence, covering 22 countries in 11 GBD world regions. No studies reporting on complete remission (equivalent to cure rather than a temporary reduction in symptom levels as clinicians tend to define ‘remission’) were found.

Prevalence and Incidence age of onset

While there is evidence to suggest that bipolar disorder commonly starts in the mid-teens or early twenties, there is still disagreement over a minimum age of onset {Goodwin, 2008 #208}. Even though symptoms can be tracked back to childhood, setting a threshold for diagnosis is difficult given that current diagnostic criteria are based on adult presentation of the disorder. The literature and expert advice on this issue suggests that although pre-pubertal bipolar disorder is rare, a distinct phenotype can exist. In order to capture as much of the burden attributable to this disorder, a minimum age of onset of 10 years and a maximum of 90 years was set for prevalence and incidence.

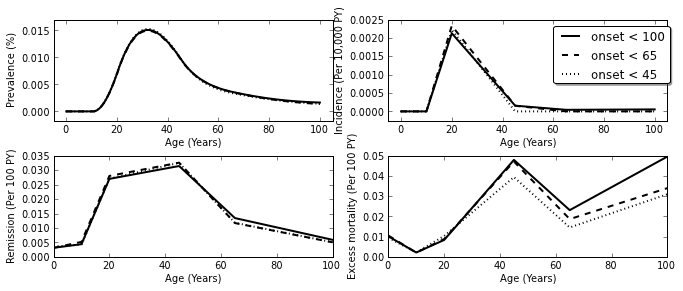
Above: Bipolar epidemiologic parameters for Western Europe males in 1990. Priors set restrictions of the age of onset so that prevalence and incidence is 0 before age 10 and after age 65.

While expert priors are useful in guiding the modeling process, they may have unintended effects as discussed in Chapter **5.2**. Choosing to have no age restrictions on incidence and prevalence, the age-specific burden of disease differs greatly, as shown below.



Above: The effect of priors can have a large effect on the disease model. Here, the prevalence of bipolar disorder in Western Europe males in 1990 differs greatly when there priors to set limits on prevalence and incidence to be 0 before age 10 and after age 90.

Like the age of onset, little is known about the upper age limit of bipolar disorder. Therefore a prior restricts the upper age limit to 65 years for incidence as it led to the most plausible fit to the data. Using expert knowledge set plausible bounds on the level of disease is useful in modeling noisy data, but changes the upper age limit produce unexpected changes as shown below.

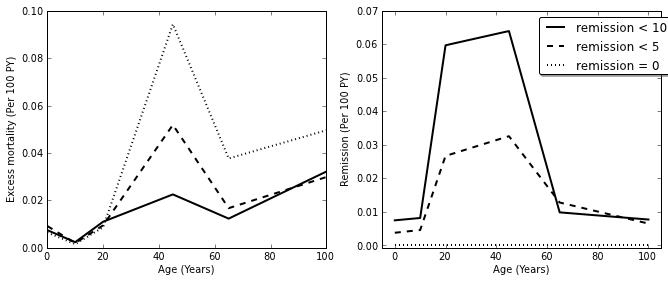


Above: Estimated prevalence, incidence, remission and excess mortality for Western Europe males with bipolar disorder in 1990. Priors that restrict the upper age limit of incidence to 45, 65 and 100 propagate through the model to changes in remission and excess mortality.

Residual v Remission

Although residual symptoms can be less severe than manic and depressive episodes, they still lead to some disability and therefore burden. However since an average episode of bipolar disorder is believed to last for about 3 months or more {Angst, 2000 #196;Goodwin, 2008 #208}, estimates of point prevalence assessing symptoms within the past month or less, will likely miss cases of bipolar disorder in a residual episode and underestimate prevalence.

The terms ‘residual’ and ‘remission’ have very different implications for GBD. A residual state involves mild symptoms with mild disability which still contribute to burden. Remission is equivalent to cure rather than a temporary reduction in symptom levels hence does not contribute to burden. Since there is no consistent use of these terms in the bipolar literature, we were unable to include any remission data in the bipolar modeling. Instead, expert guidance was sought to set an upper bound on remission hazard of 5 per 100 person-years, yielding a plausible fit to the existing data. To understand the sensitivity of model estimates to this assumption, we also considered upper bounds on remission hazard of 0 and 10 per 100 person-years. As shown in the figure below, this leads to large changes in the estimated excess mortality.



Above: With no data for bipolar disorder remission, a prior for the level of remission can have unexpected results, as shown above for Western Europe males in 1990. Changing the remission prior levels to 0.0, 0.05 or 0.1 (units?) for bipolar disorder also changes the excess mortality.

{American Psychiatric Association, 2000 #31} (DisMod Chapter\_Bipolar\_draft 2\_AF\_050911)

[2] National Center for Biotechnology Information, U.S. National Library of Medicine. “Bipolar disorder: Manic depression; Bipolar affective disorder”. 2012. <<http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0001924/>>. Accessed 10 Apr 2012.

[3] National Institute of Mental Health. 2008 (revised). 2011 (reviewed) . “Bipolar Disorder”. <<http://www.nimh.nih.gov/health/publications/bipolar-disorder/complete-index.shtml>>. Accessed 10 Apr 2012.

[4] Mayo Clinic Staff. 2012. “Bipolar Disorder”. <<http://www.mayoclinic.com/health/bipolar-disorder/DS00356>>. Accessed 10 Apr 2012.

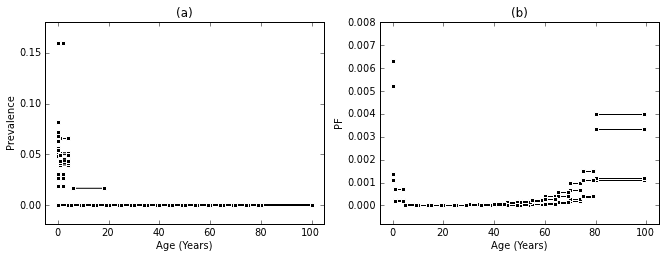
**Diarrheal Diseases**

Diarrheal diseases are a leading cause of childhood morbidity and mortality. Defined as having loose or watery stools three or more times a day or more frequently than normal for the individual, the diarrheal episodes typically last 1-8 days for acute cases [1, 2, 3]. Compared to many of the other diseases discussed thus far, diarrheal diseases have a very short duration.

Bacterial, viral and parasitic infections cause most cases of diarrhea. Non-infectious causes include drugs, surgical conditions, systemic infections and food intolerance. Typically spread via the oral-fecal route, water sanitation and hygiene is a large part of diarrhea prevention. Treatment usually involves fluid replacement, as acute diarrhea causes fluid loss and dehydration. Other common treatments are vitamin A and zinc supplementation, vaccinations and antibiotic regimes [1, 2, 3].

For the analysis of diarrheal diseases, the primary sources of data were from surveys, hospital admissions and literature review. The surveys reported period prevalence data whereas hospital admissions and the majority of the literature reported incidence data. In short term diseases, point prevalence and incidence are approximately the same. Therefore to avoid compositional bias, all incidence and period prevalence estimates were converted to point prevalence estimates using the assumption that

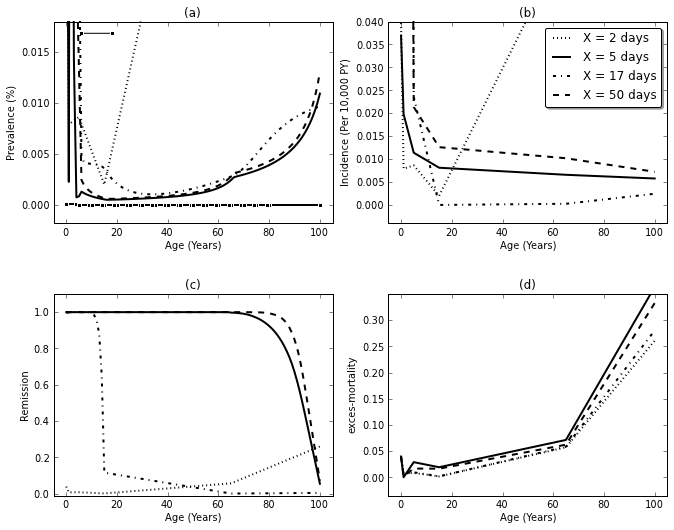
where ppoint is the point prevalence, i is the incidence and d is the duration of the disease. Thus the analysis includes 2029 prevalence estimates from 19 regions. Data without measures of uncertainty (effective sample size, standard error or 95% confidence intervals), without case definitions, and duplicated data from different sources are excluded from analysis.



Above: Diarrheal point prevalence (a) and the product of point prevalence and excess mortality (b) data in Central Latin America.

Diarrheal diseases use an integrated systems dynamic model to model all epidemiologic parameters for a single time, place and sex. An integrated systems model estimates all epidemiologic parameters simultaneously which maintains internally consistent results as discussed in Chapters **1.5** and **CKD example**. In addition to the systematic review data, informative priors on the duration of diarrheal diseases help guide the modeling process. The prior for duration is set as the prior for remission. Duration priors produce less stable estimates than the more directly parameterized remission priors. Therefore remission is the more appropriate prior for the model. The following equation approximates the relationship between duration and remission

where r is the epidemiologic rate for remission and d is the epidemiologic rate for duration in years. Similar to the discussion of priors in the **PMS example**, the choice of priors affects the epidemiological estimates. However, because of the logical requirement of internal consistency, a change in the duration of diarrheal diseases has effects on all epidemiological parameters as seen below.



Above: Estimates of prevalence (a), incidence (b), remission (c) and excess-mortality (d) for diarrheal diseases in males in Central Latin America 2005 using a consistent model. The consistent model requires internal consistency, so a change in the prior on duration has effects on all epidemiological parameters.

[1] UNICEF, WHO. 2009. “Diarrhoea: Why children are still dying and what can be done”.

[2] Carlos, CC and Saniel, MC. 1990. “Etiology and Epidemiology of Diarrhea”. *Phillips J Microbio Infect Dis*. 19(2):51-53.

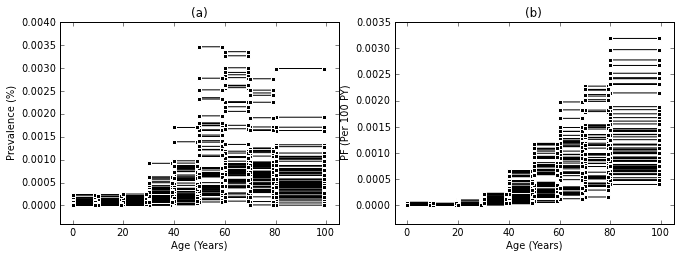
[3] Lamerti, LM, et. al. 2012. “Systematic review of diarrhea duration and severity in children and adults in low- and middle-income countries”. *BMC Public Health*. 12:276

**Cirrhosis**

Fixed effects can also increase the accuracy of out-of-sample predictions. By modeling the relationship between the epidemiologic parameter of interest and the explanatory covariate, extrapolation can predict for epidemiologic parameters in areas where no epidemiologic parameter measurements are available but covariate data are. Few regions have epidemiologic measurements for cirrhosis. However, by using the age-standardized hepatitis C prevalence as a country-level covariate to predict out-of-sample, it is possible to estimate the prevalence of cirrhosis.

The result of chronic liver injury, cirrhosis is an advanced stage of liver scarring. Cirrhosis is the end stage of any chronic liver disease, with the most common causes being alcoholic liver disease and hepatitis B and C. Asymptomatic during the early stages of the disease, compensated cirrhosis may go undetected until complications develop. The diagnostic gold standard for cirrhosis is a liver biopsy. Complications such as portal hypertension, jaundice, ascites, gastrointestinal bleeding and liver dysfunction mark the progression from compensated to decompensated cirrhosis. Irreversible, cirrhosis management is the prevention, control and treatment of cirrhosis complications, with liver transplantation being the ultimate treatment. Without a liver transplant, cirrhosis mortality is very high [1, 2, 3].

Systematic review yielded prevalence, excess mortality rate estimates from four (of 21 possible) regions. Given the difficulty in cirrhosis diagnosis, it is assumed this data represents decompensated cirrhosis and the following analysis focuses on the decompensated phase of the disease, assuming almost all prevalent cases lead to death.



Above: Available global data for cirrhosis prevalence and **pf**.

borrow strength from the mortality estimates to inform the incidence estimates

using estimates of age-standardized hepatitis C prevalence as an explanatory covariate for estimating the prevalence of cirrhosis….??

[1] Garcia-Tsao, G, et. al. 2009. “Management and Treatment of Patients With Cirrhosis and Portal Hypertension: Recommendations From the Department of Veterans Affairs Hepatitis C Resource Center Program and the National Hepatitis C Program”. *Am J Gastroenterol*. 104:1802-1829.

[2] D’Amico, G, et. al. 2006. “Natural history and prognostic indicators of survival in cirrhosis: A systematic review of 118 studies”. *Journal of Hepatology*. 44(1):217-231.

[3] Schuppan, D and Afdhal, NH. 2008. “Liver cirrhosis”. *The* *Lancet*. 371(9615):838-851.

Transport Injuries

So injuries…

know incidence and need prevalence…

more consistent models

For the 2010 GBD Study, injuries include all conditions that are codable to the ICD-9 and ICD-10 injuries chapter. This example solely focuses on transport injuries. Transport injuries include road injuries for pedestrian, bicyclist, motorized two-wheeler rider, occupation motorized vehicle with 3 or more wheels, and other road injury. It also includes other transport injury and unintentional other transport injury.

Injury is a condition codable to the ICD-9 and ICD-10 chapters with that name

only cases warranting hospitalization and cases warranting treatment by health care professional but not hospitalization included

includes short term and long term outcomes yields total global burden of injury

Transport injuries include road injury for pedestrian, bicyclist, motorized two-wheeler rider, occupation motorized vehicle with 3 or more wheels, and other road injury. It also includes other transport injury and unintentional other transport injury

modeled mortality rates from VA, surveillance systems, survey/census, police reports for mortality data

DisMod:

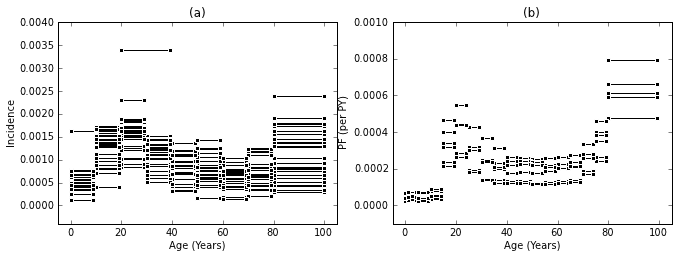
hospital records -> incidence data for cause of injury incidence

discharge databases -> proportion cases for nature of incidence -> incidence for nature of injury

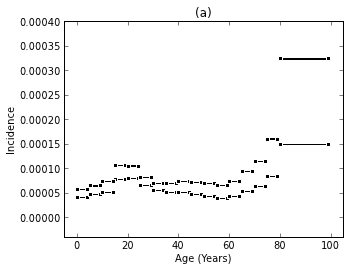
all short-term incidence -> short-term prevalence by cause or injury

portion of long-term incidence -> nature of injury prevalence

Injury incidence includes injury warranting hospital admission and injury warranting other health care. Only include a portion of incidence in calculation of prevalence. Assumes no remission and uses relative-risk of mortality and standardized mortality ration from literature reviews.



Above: Hospital and survey data for road injury and unintentional other transport injury for males in the region of North America, High Income. Systematic review yielded only incidence (a) and the product of prevalence and excess mortality (b) data.



Above: Long-term incidence data for moderate and severe traumatic brain injury in the region of North America, High Income.

Incidence data for moderate or severe traumatic brain injury warranting hospital admission or other health care in North America .

As discussed in Chapter **1.5**, epidemiologic parameters, such as incidence, prevalence, remission, and with-condition mortality, are related by a logical requirement of internal consistency. A prevalent case can only exist if there was a past incidence event and the current number of prevalence cases can be determined from past prevalent cases, new incident cases, deaths and remissions. Modeling the parameters simultaneously produces a best estimate and plausible uncertainty bounds for incidence and prevalence that are internally consistent estimates for a single time, place and sex.

Moderate or severe traumatic brain injury sequelae SS183 SS212

SS183-Moderate long-term consequences of traumatic brain injury (with or without treatment)

SS212- Severe traumatic brain injury (short term) (with or without treatment)