

# Working Specifications for Initial Measurements Under ESHB 1774 and SSHB 1566

Joseph Mienko<sup>1</sup>

<sup>1</sup>Partners for Our Children, School of Social Work  
University of Washington, Seattle, Washington

## 1. Safety

### 1.1. Safety Concerns

#### 1.1.1. General Rate of Referrals ( $R_r$ )

The general rate of referrals (i.e. reports of maltreatment) ( $R_r$ ) shall be defined as the total number of household referrals to Children's Administration (CA) per month, per 1,000 households with own children under the age of 18 as defined by the US Census. This measure will focus only on those referrals alleging child maltreatment or an eminent risk of serious harm to the child. This rate shall be calculated simply by

$$R_r = (I_r \div N_h^{<18}) \cdot 1,000 \quad (1)$$

where  $I_r$  represents the total number of first referrals occurring during the month and  $N_h^{<18}$  the population of households with own children under the age of 18 as defined by the US Census. If more than one referral is received for a given household within a 48 hour period, those referrals will only be counted once. To the extent that estimates are available, values of  $N_h^{<18}$  (and subsequent measures of the general population used in this document) shall be calculated from the following sources in order of priority:

1. US Census,
2. American Community Survey (ACS),
3. Office of Financial Management (OFM), and
4. Linear interpolation taking the above estimates as "True" values.

#### 1.1.2. General Rate of Findings ( $R_f$ )

The rate of household findings ( $R_f$ ) shall be defined as the monthly total number of first findings of child maltreatment per 1,000 referrals to CA. This measure will focus only on those referrals alleging child maltreatment or an eminent risk of serious harm to the child and will be calculated at the household level. A household is considered to have a finding if any allegation from a given referral is founded. This rate shall be calculated as

$$R_f = (I_f \div N_r) \cdot 1,000 \quad (2)$$

where  $I_f$  represents the total findings occurring during the month among referrals to the child welfare system over the same month. The value of  $N_r$  shall be calculated similarly to the value of  $I_r$  as described above.

## 1.2. Recurrence of Safety Concerns

### 1.2.1. Order-Specific Rate of Referrals ( ${}^iR_r$ )

In measuring referrals, it is important to distinguish between first referrals and subsequent referrals. In general, those households that experience at least one referral to the child welfare system are at an increased risk of rereferral. As such, there is interest in distinguishing between first and higher order referrals (e.g. second referrals, third referrals, etc.). However, the population at risk for higher-order referrals is different for each order.

The order-specific rereferral rate is defined as the number of referrals of a given order during a month per 1,000 households with own children under the age of 18. The general formula for  $i$ th order rereferral rates is

$${}^iR_r = \frac{{}^iI_r}{{}_{pr}^{i-1}N_h^{<18}} \cdot 1,000 \quad (3)$$

where  ${}^iI_r$  represents the total number of  $i$ th order referrals and  ${}_{pr}^{i-1}N_h^{<18}$  represents the population of households with  $i$ th – 1 order referrals in which all children in the household ( $h$ ) are still under the age of 18. At this time, it is proposed that only first and second-order rereferrals be used as measures under 1774. Future reports may include higher-order rereferrals.

### 1.2.2. Order-Specific Rate of Findings ( ${}^iR_f$ )

Similarly to referrals, those households that experience at least one finding of maltreatment to the child welfare system are at an increased risk of findings of maltreatment. As such, there is interest in distinguishing between first and higher order findings (e.g. second findings, third findings, etc.). However, the population at risk for higher-order findings is different for each order.

The order-specific finding rate is defined as the number of findings of a given order during a month per 1,000 families with referrals who have received  $i - 1$  prior findings. The general formula for  $i$ th order rereferral rates is

$${}^iR_f = \frac{{}^iI_f}{{}_{pf}^{i-1}N_r} \cdot 1,000 \quad (4)$$

where  ${}^iI_f$  represents the total number of  $i$ th order findings and  ${}_{pf}^{i-1}N_r^{<18}$  represents the population of referrals with  $i$ th – 1 order findings. As with order-specific referrals, it is proposed that only first and second-order refindings be used as measures under 1774. Future reports may include higher-order findings.

## 1.3. Placement in Out-Of-Home Care

### 1.3.1. General Rate of Placement ( $R_p$ )

The general rate of placement into out-of-home care ( $R_p$ ) shall be defined as the monthly total number of placements to Children’s Administration per 1,000 referrals to CA. This rate shall be calculated simply by

$$R_p = (I_p \div N_r) \cdot 1,000 \quad (5)$$

where  $I_p$  represents the total number of placements occurring during the month and  $N_r$  the number of referrals (as defined above).

### 1.3.2. Order-Specific Placement Rate ( ${}^iR_p$ )

Similarly to referrals and findings, those households that experience at least one placement in out-of-home care are at an increased risk of placement during the investigation of subsequent referrals. As such, there is interest in distinguishing between first and higher order placements (e.g. second placements, third placements, etc.). However, the population at risk for higher-order placements is different for each order.

The order-specific placement rate is defined as the number of placements of a given order during a month per 1,000 referrals who have received  $i - 1$  prior placements. The general formula for  $i$ th order placement rates is

$${}^iR_p = \frac{{}^iI_p}{{}^{i-1}N_{r_{pp}}} \cdot 1,000 \quad (6)$$

where  ${}^iI_f$  represents the total number of  $i$ th order placements and  ${}^{i-1}N_{r_{pf}}^{<18}$  represents the population of referrals with  $i$ th - 1 order placements. As with order-specific referrals and findings, it is proposed that only first and second-order placements be used as measures under 1774. Future reports may include higher-order placements.

## 1.4. Maltreatment in Foster Care

### 1.4.1. Care-Day Finding Rate ( ${}_dR_f$ )

The final safety measurement concerns the rate of findings per day of foster care. This is the first “care-day” measurement proposed in this document - measurements in which the population of interest is the is a count of days ( $d$ ) of out-of-home care for some population as opposed to a count of some type of person.

The care-day finding rate follows the similarly calculated proposed CFSR measurement where  ${}_dR_f$  is defined as

$${}_dR_f = ({}_dI_f \div N_d) \cdot 100,000 \quad (7)$$

where  $N_d$  represents the total number of care days provided by CA in a given year and  ${}_dI_f$  represents the number of maltreatment findings for children taking place during a day of care over the same period. This measurement excludes children in foster care for less than 8 days and also excludes reports made within the first 7 days of the referral.

## 2. Permanency

### 2.1. Placement Mobility

#### 2.1.1. Duration-Specific Movement Rate ( $R_m^j$ )

This measurement will adopt the basic approach of the analagous proposed CFSR measurement in which total placements are measured against the number of care days experienced for children during their first year of care. In this measurement, we propose that we

extend this basic logic to a duration-specific mobility measurement in which a care-day mobility rate is calculated for the first through tenth duration periods ( $j$ ) of care such that  $R_m^j$  is given as

$$R_m^j = (I_m^j \div N_d) \cdot 100,000 \quad (8)$$

where  $I_m^j$  is the count of placement moves during the  $j$ th year and  $C^j$  is the number of care days associated with the  $j$ th year of a particular episode of care. Care days associated with children who stay in care for less than 8 days in their original placement episode will be excluded from  $N_d$  in this measurement and moves to respite or other temporary situations will be excluded from  $I_m^j$  in this measure.

### 2.1.2. Transition-Duration-Specific Mobility Rate ( $_{tx}R_m^j$ )

In order to get a sense of the *types* of transitions that are experienced by children,  $R_m^j$  shall be extended to also measure the type of transition that children are moving toward ( $tx$ ) (i.e. foster care ( $f$ ), relative placement ( $r$ ), and group care ( $g$ )). The  $_{tx}R_m^j$  measurement shall be calculated identically to  $R_m^j$  above. However,  $_{tx}R_m^j$  shall be calculated separately for  $tx = f$ ,  $tx = r$ , and  $tx = g$ .

### 2.1.3. Age-Order-Specific AWOL Rate ( $^iR_{awol}^a$ )

This measurement will extend the care-day approach to the measurement of how well the agency is avoiding children on AWOL status. This measurement will measure the number of AWOL care days per 100,000 care days within a given year.

$$^iR_{awol}^a = \frac{^i_{awol}I_d^a}{N_d^{\geq 9}} \cdot 100,000 \quad (9)$$

where  $^i_{awol}I_d^a$  is the number of  $i$ th order AWOL care-days for children of age  $a$  and  $N_d^{\geq 9}$  is the number of care days (restricted similarly to the above measures) associated with children where  $a \geq 9$ .

## 2.2. Time to Permanence

### 2.2.1. Percentage of Children Experiencing Permanency Before 1 Year ( $P_{pm}^1$ )

This measurement is similar to the proposed CFSR measurement requiring exit to permanency within 12 months. However, instead of calculating a specific percentage, we propose that a Kaplan-Meier estimator be used to better account for right censoring of children who either age-out of the system or exit to a form of permanency other than adoption, reunification, or some form of guardianship.

The Kaplan-Meier estimator for a federally recognized form of permanency  $pm$  is given as

$$_{pm}\hat{S}(t) = \prod_{t_i < t} \frac{n_i - pm_i}{n_i} \quad (10)$$

where  $n_i$  is the number of children entering out-of-home care in a given fiscal year who are still in care just prior to time  $t_i$ , and  $pm_i$  is the number of permanency events at  $t_i$ .

Thus, the probability of exit within 1 year is given as  ${}_{pm}\hat{S}(365)$ . A reasonable approximation of the new CFSR measurement is given as  $(1 - {}_{pm}\hat{S}(365)) \cdot 100$  and is thus the proposed definition of this measure. All children who stay in care for less than 8 days will be excluded from this measurement. Children aging out of care or experiencing a permanency outcome other than reunification, adoption, or some form of guardianship will be treated as right-censored to observation.

### 2.2.2. Percentage of Children Experiencing Adoption Before 1 Year ( $P_{ad}^1$ )

This measurement will also make use of a Kaplan-Meier estimator. Specifically, the Kaplan-Meier estimator for adoption  $ad$  is given as

$${}_{ad}\hat{S}(t) = \prod_{t_i < t} \frac{n_i - ad_i}{n_i} \quad (11)$$

where  $n_i$  is the number of children entering legally-free status in a given fiscal year who have not yet been adopted just prior to time  $t_i$ , and  $ad_i$  is the number of adoption events at  $t_i$ . Thus, the probability of adoption within 1 year is given as  ${}_{ad}\hat{S}(365)$ . Children aging out of care or experiencing a permanency outcome other than adoption will be treated as right-censored to observation. For consistency with the proposed CFSR permanency measure,  $(1 - {}_{ad}\hat{S}(365)) \cdot 100$  is the proposed definition of this measure.

### 2.2.3. Care-Day Rate ( $R_d$ )

This measurement will further extend the use of care-days - this time as an incidence as opposed to a population of interest. Specifically, this measure will report the rate of care-days per exit from care. This measure shall be calculated simply as

$$R_d = (I_d \div N_{pm}) \cdot 100,000 \quad (12)$$

where  $I_d$  is the number of care-days generated over a given reporting period and  $N_{pm}$  is the number of federally recognized forms of permanency ( $pm$ ) as described above.

## 2.3. Duration of Permanence

### 2.3.1. Percentage of Children Reentering Care Within 1 Year ( ${}_R\hat{S}(365)$ )

This measurement will also make use of a Kaplan-Meier estimator. Specifically, the Kaplan-Meier estimator for reentry  $rt$  is given as

$${}_{rt}\hat{S}(t) = \prod_{t_i < t} \frac{n_i - rt_i}{n_i} \quad (13)$$

where  $n_i$  is the number of children entering permanency in a given fiscal year who have not yet reentered care just prior to time  $t_i$ , and  $rt_i$  is the number of reentry events at  $t_i$ . Thus, the probability of reentry within 1 year is given as  ${}_{rt}\hat{S}(365)$ . Children aging out of care or experiencing a permanency outcome other than adoption will be treated as right-censored to observation. Children who stay in care for less than 8 days will be excluded from this measurement. For consistency with the proposed CFSR permanency measure,  $(1 - {}_{rt}\hat{S}(365)) \cdot 100$  is the proposed definition of this measure.

### 3. Well-Being

#### 3.1. Maintenance of Family Relationships

##### 3.1.1. Sibling Placement Rate ( $R_{sp}$ )

This measurement will also utilize a care-day approach to report the extent to which agency care-days tend to involve the joint placement of sibling groups. Specifically, this measurement will report the number of care-days where siblings are placed together per 100,000 sibling care-days. The measurement ( $R_{sp}$ ) is specifically defined as

$$R_{sp} = (tgI_{ds} \div N_{ds}) \cdot 100,000 \quad (14)$$

where  $N_{ds}$  is the number of days of care associated with children that are members of a recorded sibling group and  $tgI_{ds}$  is the number of days of care associated with children that are members of a recorded sibling group, where the sibling group was placed together. Restrictions will be placed on  $N_{ds}$  similarly to other measurements.

#### 3.2. Educational Readiness

##### 3.2.1. Post-Secondary Planning Rate ( $R_{psp}$ )

The post-secondary planning rate  $R_{psp}$  shall be defined as the total number of children per fiscal year exiting with some record of post-secondary planning activity. This rate shall be calculated simply by

$$R_{psp} = (I_{e_{psp} \geq 17} \div N_e^{\geq 17}) \cdot 100,000 \quad (15)$$

where  $I_{e_{psp} \geq 17}$  is defined as the number of children exiting where the child's age is  $\geq 17$  with some record of the child being provided with assistance in completing college applications, obtaining financial aid, planning for college tours, or some other form of post-secondary activity and  $N_e^{\geq 17}$  is defined as the number of children exiting care where the child's age is  $\geq 17$ . This measure will be limited to children in care for more than 7 days.

##### 3.2.2. Age-Specific-Kindergarten Readiness Rate ( $P_{kr}$ )

The age-specific-kindergarten readiness rate shall be defined as the ratio of kindergarten readiness among kindergarten foster children, expressed as a percentage. The percentage shall be calculated as

$$P_{kr} = (I_{kr}^a \div N_K^a) \cdot 100$$

where  $I_{kr}^a$  is the age-specific number of foster children in Kindergarten performing at (or above) expected levels across all six developmental areas as assessed by the WaKIDS assessment for a particular academic year and  $N_K^a$  is the age-specific number of foster children in kindergarten during a particular academic year. Children shall be excluded from this measurement if they did not complete the WaKIDS assessment.

### 3.2.3. Age-Specific-Third Grade Literacy Rate ( $P_{kr}^a$ )

The age-specific-third-grade literacy rate shall be defined as the ratio of third-grade reading proficiency among third grade foster children, expressed as a percentage. The percentage shall be calculated as

$$P_{3l}^a = (I_{3l}^a \div N_3^a) \cdot 100$$

where  $I_{3l}^a$  is the age-specific number of foster children in third-grade performing at (or above) expected levels as measured by the MSP for a particular academic year. For this measure, children should be excluded from  $N_3^a$  if they did not complete the MSP assessment.

## 3.3. Adult Functioning

### 3.3.1. Grade-Age-Specific-Cumulative Grade Attainment Rate ( $P_{g+}^a$ )

The grade-age-specific cumulative grade attainment rate shall be defined as the rate of grade completion per former foster child for a given grade and age. This measurement shall be expressed as a percentage and calculated as

$$\frac{{}_{y+n}C_a^{g+}}{{}_yP_{17}^g}$$

where  ${}_yP_{17}^g$  is the count of persons in a grade-specific cohort of 17 year-old foster children (i.e. foster children who will reach the age-of-majority prior to the start of the following academic year) and  ${}_{y+n}C_a^{g+}$  is the count of persons in a given cohort who have completed grade  $g$  or beyond by the end of  $n$  academic years (including  $n = 0$ , the current academic year). Using a combination of data from PCHEES, NSC, and CEDARS, grade attainment shall be reported based on the follow-up time available. In general, the extent to which a particular value of  $g^+$  will be reported shall be determined as a function of  $n$  as shown in the table below. The categories of  $g^+$  are based on US Census categories of educational [attainment](<https://www.census.gov/hhes/socdemo/education/about/>).

### 3.3.2. Incidence of reports of safety concerns

#### The incidence of reports of household safety concerns

**Definition** The incidence of reports of household safety concerns ( $IR_R$ ) shall be defined as the total number of first referrals to Children's Administration (CA) per month, per 1,000 households with own children under the age of 18 as defined by the US Census. This measure will focus only on those referrals alleging child maltreatment or an eminent risk of serious harm to the child. This rate shall be calculated simply by

$$IR_R = (I_R \div P_{H<18}) \cdot 1,000 \quad (16)$$

where  $I_R$  represents the total number of first referrals occurring during the month and  $P_{H<18}$  the number of households with own children under the age of 18 as defined by the US census. To the extent that estimates are available, values of  $P_{H<18}$  shall be calculated from the following sources in order of priority:

1. US Census,
2. American Community Survey (ACS),
3. Office of Financial Management (OFM), and
4. Linear interpolation taking the above estimates as “True” values.

**SQL Code** The SQL code used to obtain data for this measurement is as follows:

```
declare @fystart int
declare @fystop int

set @fystart = 2010
set @fystop = 2013;

select
    count(*) I_R
    ,P_H
    ,count(*)*1000.0/P_H IR_R
    ,month calendar_month
from ca_ods.dbo.calendar_dim cd
join ca_ods.base.tbl_intakes ti
    on ti.inv_ass_start = cd.calendar_date
join (select
    measurement_year
    ,sum(pop_cnt) P_H
    from ca_ods.dbo.ref_lookup_census_population_poc2
    where cd_race < 8
    group by
        measurement_year) pop
    on pop.measurement_year = year(cd.calendar_date)
where state_fiscal_yyyy between @fystart and @fystop
    and cd_access_type in (1, 4)
    and intake_rank = 1
group by
    month
    ,P_H
```

As can be seen, we make use of a previously calculated base table named `tbl_intakes`. This table is based on a select FROM the the CA reporting layer table named `rptIntakes` and a LEFT JOIN to `INTAKE.FACT` from `FLDW`. As such, the initial universe of referrals is identical to those referrals (i.e. intakes) utilized by CA for reporting purposes.

In addition to `tbl_intakes`, we use a previously calculated table of interpolated values  $P_H$  (`ref_lookup_census_population_poc2`) and the (`calendar_dim`) from `FLDW`. All code is included here for the sake of transparency.



**R Code** The first chunk of code simply clears our memory, sets the working directory, loads the RODB library (an R library which enables an analyst to connect R to an ODBC database), and loads the ggplot2 library (an R library for plotting). We also assign a vector of visually pleasing colors (`poc_colors`) and define a custom function (`read.sql`) which will allow us to read the text of our SQL scripts into RODB functions.

```
rm(list=ls()) ###Clear Memory
setwd("C:/Users/mienkoja/Dropbox/repos/1774")
require(RODBC) ###Load library

## Loading required package: RODBC

require(ggplot2)

## Loading required package: ggplot2

poc_colors <- c("#3B6E8F", "#A2B69A", "#A3DCE6", "#A784B4")

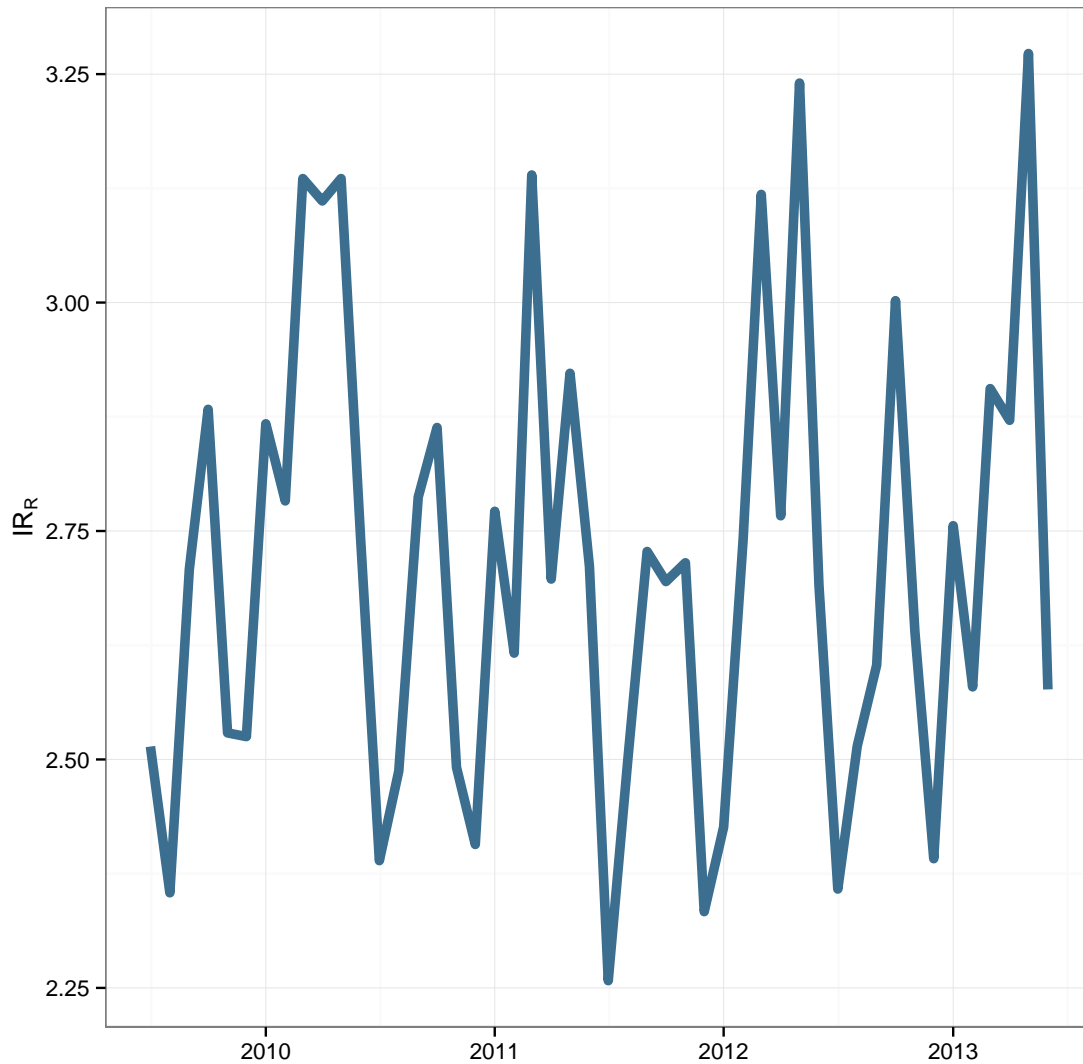
# function to read in sql files
read.sql <- function(filename, silent = TRUE) {
  q <- readLines(filename, warn = !silent)
  # remove full-line comments
  q <- q[!grepl(pattern = "^\\s*--", x = q)]
  # remove midline comments
  q <- sub(pattern = "--.*", replacement="", x = q)
  q <- paste(q, collapse = " ")
  return(q)
}

# Assign database connection
cn <- odbcConnect("POC")
```

In this next chunk of code, I read the sql script outlined above into a dataframe named (`dat.IR_R`). We then plot the monthly time series of  $I_R$  as shown in the figure below.

```
dat.IR_R <- sqlQuery(cn, read.sql("IR_R.sql"))
dat.IR_R$calendar_month <- as.Date(dat.IR_R$calendar_month)

#plot time series
ggplot(dat.IR_R, aes(y=IR_R, x=calendar_month)) +
  geom_line(colour=poc_colors[1], size=2) +
  labs(y = expression(IR[R]), x = "") +
  theme_bw()
```



### The incidence of placement into out-of-home care

**Definition** The incidence of placement into out-of-home care ( $IR_P$ ) shall be defined as the monthly total number of referrals to Children’s Administration per 1,000 persons under the age of 18. This rate shall be calculated simply by

$$IR_P = (I_P \div P_C) \cdot 1,000 \quad (17)$$

where  $I_P$  represents the total number of placements occurring during the month and  $P_C$  the number of children under the age of 18 in the general population. To the extent that estimates are available, values of  $P_C$  shall be calculated from the same sources and order of priority as in  $P_H$  above.

**SQL Code** In this section I use a previously calculated table of interpolated values  $P_C$  (`ref.lookup_census_population`), a CA “reporting layer” table of place-

ments into out-of-home care (rptPlacement), and the aforementioned (calendar\_dim) table.

```
declare @fystart int
declare @fystop int

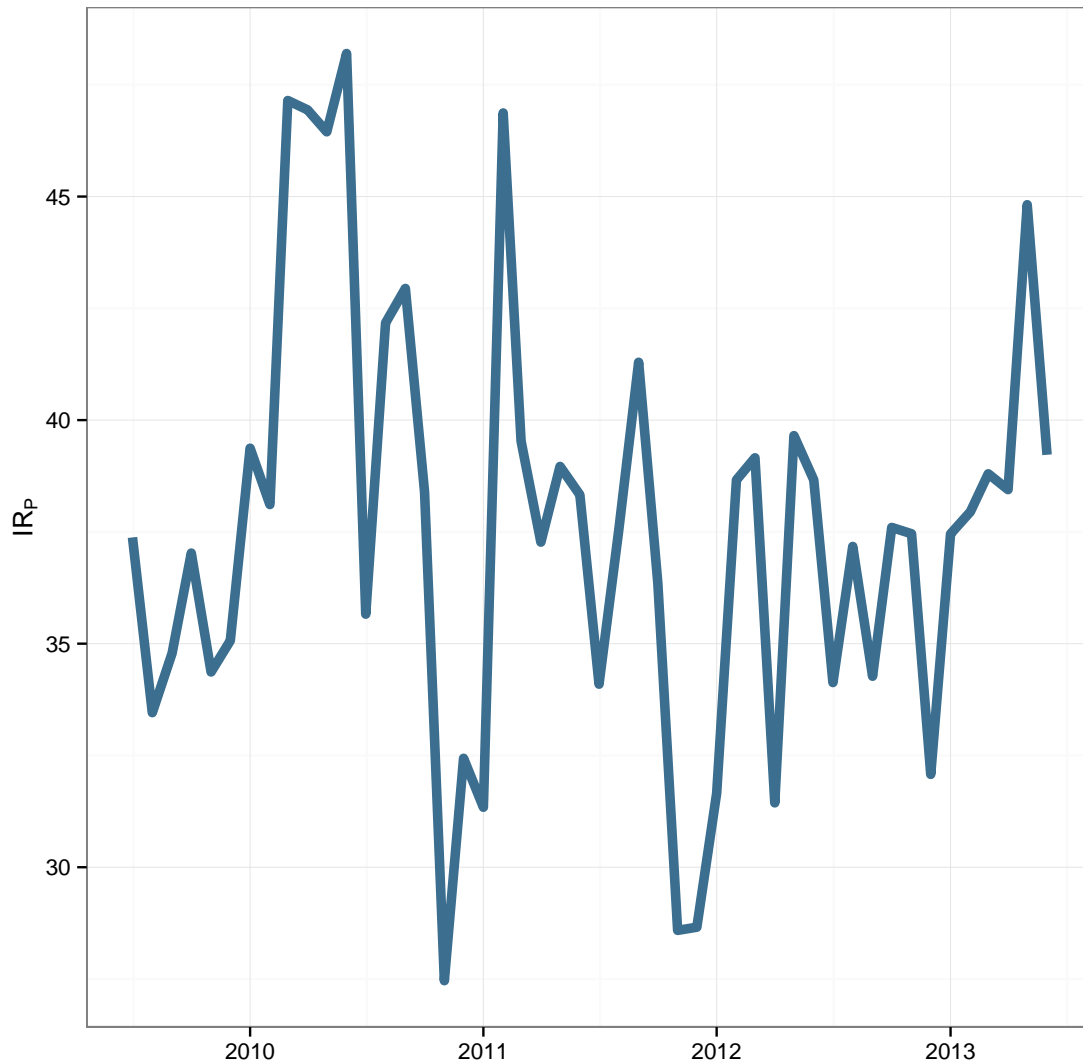
set @fystart = 2010
set @fystop = 2013;

select
  count(*) I_P
  ,P_C
  ,count(*)*100000.00/P_C IR_P
  ,month
from ca_ods.dbo.calendar_dim cd
join ca_ods.base.rptPlacement rp
  on rp.removal_dt = cd.calendar_date
join (select
  measurement_year
  ,sum(pop_cnt) P_C
  from ca_ods.dbo.ref_lookup_census_population
  where cd_race < 8
  group by
    measurement_year) pop
  on pop.measurement_year = year(cd.month)
where state_fiscal_yyyy between @fystart and @fystop
group by
  month
  ,P_C
order by
  month
```

**R Code** We plot the monthly time-series in a manner similar to what we did with  $IR_R$ .

```
dat.IR_P <- sqlQuery(cn, read.sql("IR_P.sql"))

#plot time series
ggplot(dat.IR_P, aes(y=IR_P, x=month)) +
  geom_line(colour=poc_colors[1], size=2) +
  labs(y = expression(IR[P]), x = "") +
  theme_bw()
```



## Placement rate

**Definition** The rate of placement ( $PR$ ) shall be defined as the monthly total number of first placements per 1,000 referrals to CA with no prior history of a placement in the household. For consistency with previous measurements, this measure will focus only on those referrals alleging child maltreatment or an eminent risk of serious harm to the child and will be calculated at the household level. A household is considered to have a placement if any child associated with the home on a given referral is placed into out-of-home care. This rate shall be calculated as

$$PR = (P \div I_{R_{NPP}}) \cdot 1,000 \quad (18)$$

where  $P$  represents the total placements occurring during the month among referrals for households with no previous placement ( $I_{R_{NPP}}$ ).

**SQL Code** The following SQL code outlines the continued use of (tbl\_intakes). We make use of additional fields beyond the reporting layer fields which have been calculated according to code reviewed previously by CA.

```
declare @fystart int
declare @fystop int

set @fystart = 2010
set @fystop = 2013;

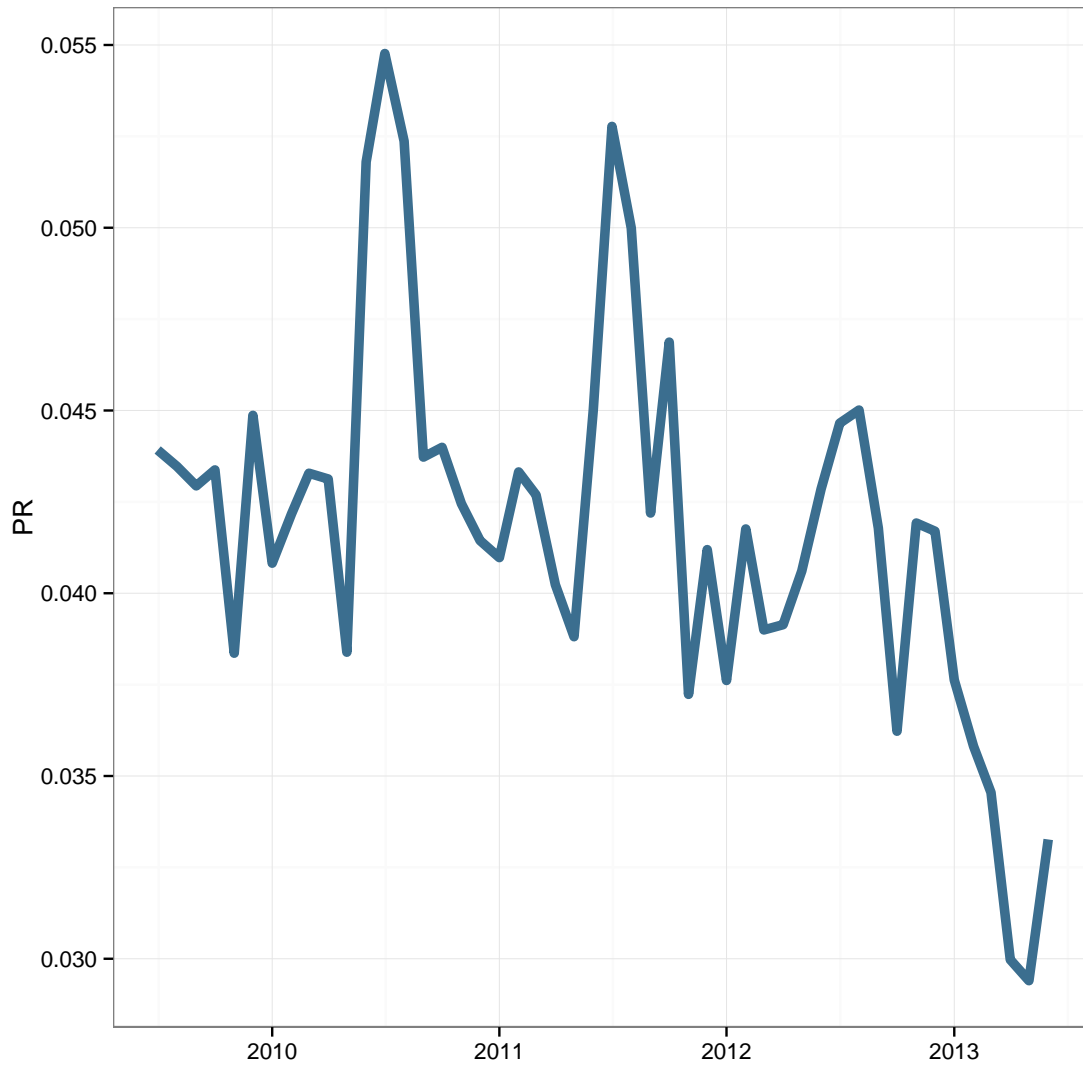
select
    count(inf.id_intake_fact) I_R1
    ,sum(iiif(inf.fl_ooh_after_this_referral=1, 1, 0)) P
    ,sum(iiif(inf.fl_ooh_after_this_referral=1, 1, 0))*1.0/
    count(inf.id_intake_fact) PR
    ,cd.month
    ,cd.state_fiscal_yyyy fiscal_yr
from ca_ods.dbo.calendar_dim cd
    left join dbCoreAdministrativeTables.dbo.tbl_intakes inf
        on cd.id_calendar_dim between
            convert(int, convert(varchar, inf.inv_ass_stop, 112))
            and convert(int, convert(varchar, inf.inv_ass_stop, 112))
where
    cd.state_fiscal_yyyy between @fystart and @fystop
    and cd_access_type in (1, 4)
    and (inf.intake_rank = 1 or fl_ooh_prior_this_referral = 0)
group by
    cd.month
    ,cd.state_fiscal_yyyy
order by
    month
```

Again, we plot the monthly time-series using R.

```
dat.PR <- sqlQuery(cn, read.sql("PR.sql"))

#plot time series
ggplot(dat.PR, aes(y=PR, x=month)) +
    geom_line(colour=poc_colors[1], size=2) +
    labs(y = expression(PR), x = "") +
    theme_bw()
```

**R Code**



## Finding rate

**Definition** The rate of household findings ( $FR$ ) shall be defined as the monthly total number of first findings of child-abuse or neglect per 1,000 referrals to CA with no prior history of a founded allegation of child abuse or neglect. This measure will focus only on those referrals alleging child maltreatment or an eminent risk of serious harm to the child and will be calculated at the household level. A household is considered to have a finding if any allegation from a given referral is founded. This rate shall be calculated as

$$FR = (F \div I_{R_{NPF}}) \cdot 1,000 \quad (19)$$

where  $F$  represents the total findings occurring during the month among referrals for households with no previous finding ( $I_{R_{NPF}}$ ).

**SQL Code** The following SQL code outlines the continued use of (tbl\_intakes). Here, we make use of additional fields beyond the reporting layer fields which have been calculated according to code reviewed previously by CA.

```
declare @fystart int
declare @fystop int

set @fystart = 2010
set @fystop = 2013;

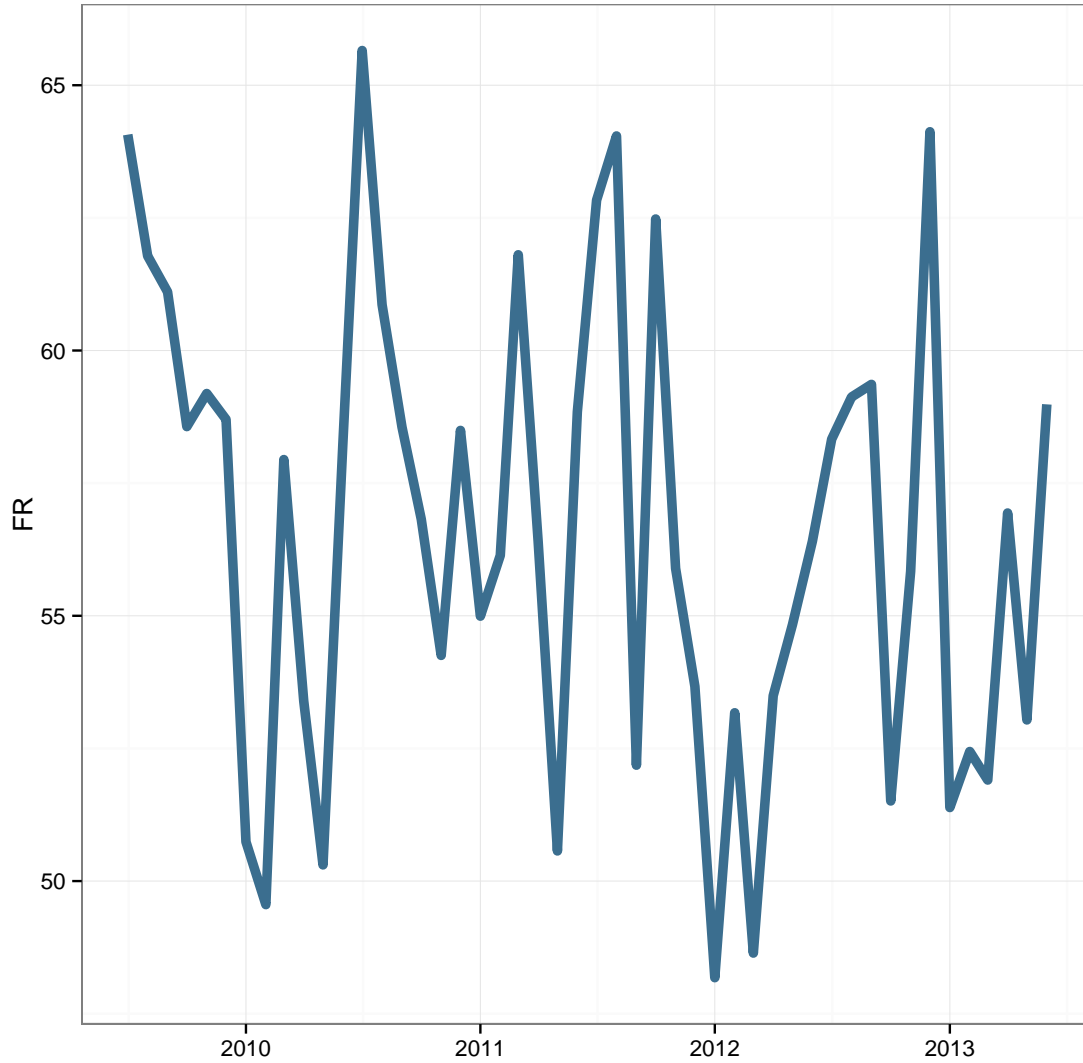
select
  count(inf.id_intake_fact) I_RNPF
  ,sum(inf.fl_founded_any_legal) F
  ,sum(inf.fl_founded_any_legal)*1000.0/
  count(inf.id_intake_fact) FR
  ,cd.month
  ,cd.state_fiscal_yyyy fiscal_yr
from ca_ods.dbo.calendar_dim cd
left join ca_ods.base.tbl_intakes inf
on cd.id_calendar_dim between
  convert(int, convert(varchar, inf.inv_ass_stop, 112))
  and convert(int, convert(varchar, inf.inv_ass_stop, 112))
where cd.state_fiscal_yyyy between @fystart and @fystop
and cd_access_type in (1, 4)
and (inf.intake_rank = 1 or fl_founded_prior_any_legal = 0)
group by
  cd.month
  ,cd.state_fiscal_yyyy
order by
  month
```

Again, we plot the monthly time-series using R.

```
dat.FR <- sqlQuery(cn, read.sql("FR.sql"))

#plot time series
ggplot(dat.FR, aes(y=FR, x=month)) +
  geom_line(colour=poc_colors[1], size=2) +
  labs(y = expression(FR), x = "") +
  theme_bw()
```

**R Code**



### 3.3.3. Recurrence of safety concerns

**Order-Specific Rereferral rate** In measuring rereferrals, it is important to distinguish between first referrals ( $IR_S$ ) and rereferrals. In general, those households that experience at least one referral to the child welfare system are at an increased risk of rereferral. As such, there is interest in distinguishing between first and higher order referrals (e.g. second referrals, third referrals, etc.).

**Definition** The order-specific rereferral rate is defined as the number of referrals of a given order during a month per 1,000 households with own children under the age of 18. The formula for the first referral is given in formula 16. The formula for  $i$ th order rereferral rates is

$$RRR_i = \frac{I_{R_i}}{P_{i-1PR<18}} \cdot 1,000 \quad (20)$$



where  $I_{R_i}$  represents the total number of  $i$ th order rereferrals and  $P_{i-1PR<18}$  represents the number of  $i$ th - 1 order referrals in which all children in the household are still under the age of 18.

**SQL Code** In the following two sections of code we demonstrate the calculation of  $RRR_2$  or second-order re-referrals. At this time, it is proposed that only second-order re-referrals be used as measures under 1774. Future reports may include higher-order re-referrals. In order to perform this calculation we make use of our base working table of intakes (tbl\_intakes), the (calendar\_dim), and a table of all children who appear to be associated with a given intake (tbl\_household\_children). For the sake of transparency, we use four separate queries to complete this measurement:

1. A simple count of open intakes on the first day of 2010 (RRR\_i\_start.sql),

```
declare @fystart int

set @fystart = 2010;

with hh_w_ch as
(
select
    ti.id_intake_fact
    ,isnull(max(dateadd(yy, 18, thc.dt_birth))
        ,dateadd(yy
            ,18
            ,ti.inv_ass_start)) mx_dt_18
    ,isnull(min(thc.dt_birth)
        ,ti.inv_ass_start) mn_dt_br
from ca_ods.base.tbl_intakes ti
    left join ca_ods.base.tbl_household_children thc
        on ti.id_intake_fact = thc.id_intake_fact
        and ti.id_case = thc.id_case
where
    cd_access_type in (1, 4)
    and intake_rank = 1
group by
    ti.id_intake_fact
    ,ti.inv_ass_start
)
select
    count(distinct id_intake_fact) pop_start
from hh_w_ch
where isnull(mx_dt_18, '3999-01-01') >=
    datefromparts(@fystart-1, 07, 01)
    and mn_dt_br < datefromparts(@fystart-1, 07, 01)
```

2. Two separate selections give us new first intakes from 2010 to present (RRR\_i\_entry3.sql and RRR\_i\_entry2.sql), and

```
declare @fystart int
```

```

declare @fystop int

set @fystart = 2010
set @fystop = 2013;

select
    count(*) pop_enter
    ,month
from
    ca_ods.base.tbl_intakes ti
    left join ca_ods.dbo.calendar_dim cd
        on ti.inv_ass_start = calendar_date
where intake_rank = 1
    and cd_access_type in (1,4)
    and inv_ass_start
    between datefromparts(@fystart-1, 07, 01)
    and datefromparts(@fystart, 06, 30)
group by
    month
order by
    month

```

```

declare @fystart int
declare @fystop int

set @fystart = 2010
set @fystop = 2013;

select
    count(*) pop_enter
    ,month
from ca_ods.base.tbl_intakes ti
    left join ca_ods.dbo.calendar_dim cd
        on ti.inv_ass_start = calendar_date
where intake_rank = 2
    and cd_access_type in (1,4)
    and inv_ass_start
    between datefromparts(@fystart-1, 07, 01)
    and datefromparts(@fystop, 06, 30)
group by
    month
order by
    month

```

### 3. A selection of families who are “aging-out” of $P_{1PR<18}$ (RRR\_i\_exit.sql)

```

declare @fystart int
declare @fystop int
set @fystart = 2010
set @fystop = 2013;
with max_dt_18 as

```

```

(select
  ti.id_intake_fact
  ,isnull(max(dateadd(yy, 18, thc.dt_birth))
    ,dateadd(yy
    ,18
    ,ti.inv_ass_start)) mx_dt_18
from ca_ods.base.tbl_intakes ti
  left join ca_ods.base.tbl_household_children thc
    on ti.id_intake_fact = thc.id_intake_fact
    and ti.id_case = thc.id_case
where
  cd_access_type in (1,4)
  and intake_rank = 1
group by
  ti.id_intake_fact
  ,ti.inv_ass_start
having isnull(max(dateadd(yy, 18, thc.dt_birth))
  ,dateadd(yy
  ,18
  ,ti.inv_ass_start)) <=
  datefromparts(@fystop, 06, 30))
select
  count(distinct id_intake_fact) pop_exit
  ,month
from ca_ods.dbo.calendar_dim cd
  left join max_dt_18 m18
    on mx_dt_18
    between dateadd(mm, -1, month) and month
where month
  between datefromparts(@fystart-1, 07, 01)
  and datefromparts(@fystop, 06, 30)
group by
  month
order by
  month

```

**R Code** The following code demonstrated how we can load these four separate queries into R for further calculation:

```

dat.RRR_i.s <- sqlQuery(cn, read.sql("RRR_i_start.sql"))
dat.RRR_i.e1 <- sqlQuery(cn, read.sql("RRR_i_entry1.sql"))
dat.RRR_i.e2 <- sqlQuery(cn, read.sql("RRR_i_entry2.sql"))
dat.RRR_i.ex <- sqlQuery(cn, read.sql("RRR_i_exit.sql"))
P_1PRlt18 <- dat.RRR_i.s$pop_start +

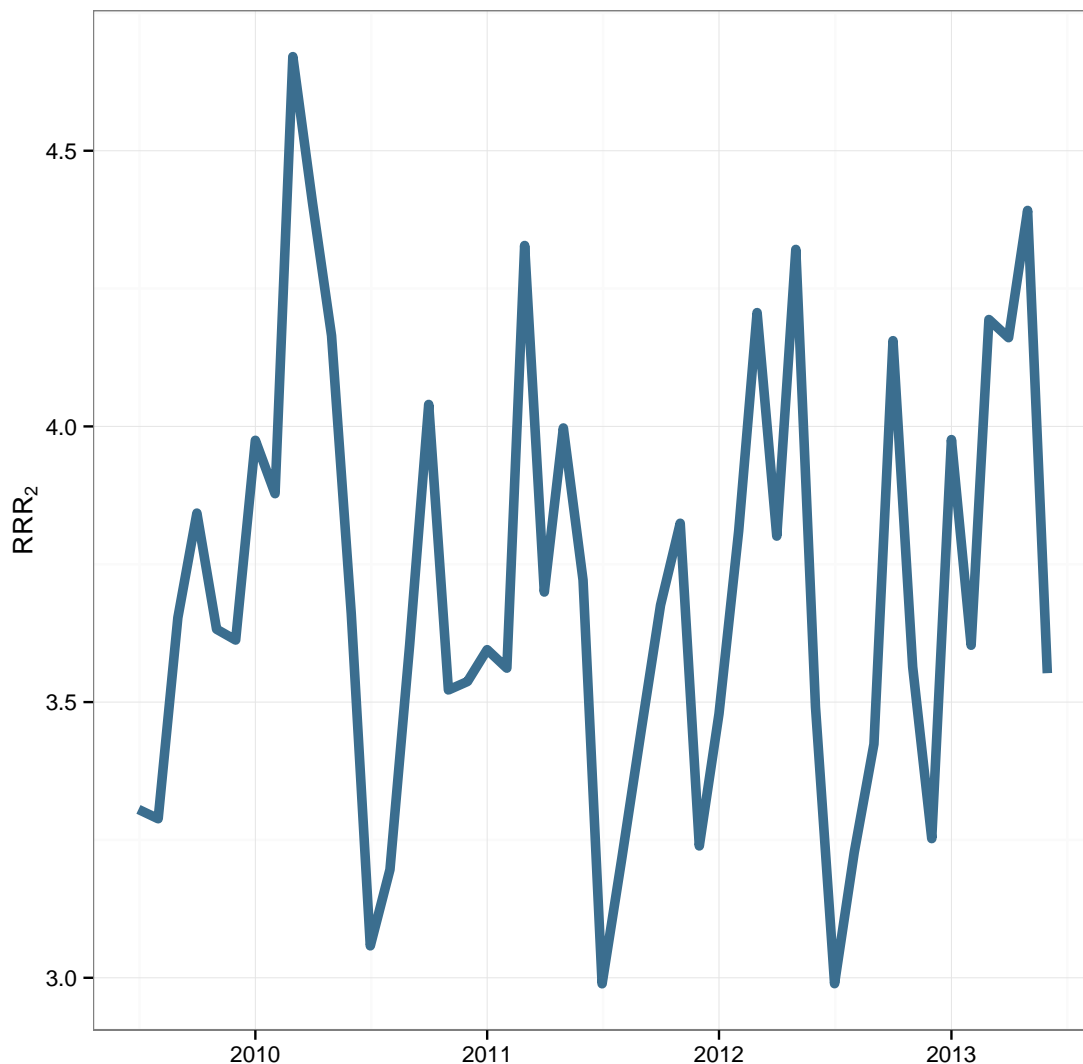
```

```

dat.RRR_i.e1$pop_enter -
dat.RRR_i.ex$pop_exit
I_R2 <- dat.RRR_i.e2$pop_enter
dat.RRR_i.e2$RRR2 <- (I_R2/P_1PR1t18)*1000

#plot time series
ggplot(dat.RRR_i.e2, aes(y=RRR2, x=month)) +
  geom_line(colour=poc_colors[1], size=2) +
  labs(y = expression(RRR2), x = "") +
  theme_bw()

```



**Order-Specific Refinding rate** Similarly to re-referrals, those households that experience at least one finding (*FR*) to the child welfare system are at an increased risk of refinding. As such, there is interest in distinguishing between first and higher order refer-

rals (e.g. second referrals, third referrals, etc.).

**Definition** The order-specific refinding rate is defined as the number of findings of a given order during a month per 1,000 households with own children under the age of 18 who have received  $i - 1$  prior referrals. The formula for the first referral is given in formula 19. The formula for  $i$ th order rereferral rates is

$$RFR_i = \frac{I_{F_i}}{P_{i-1PF<18}} \cdot 1,000 \quad (21)$$

where  $I_{F_i}$  represents the total number of  $i$ th order refindings and  $P_{i-1PF<18}$  represents the number of  $i - 1$  order findings in which all children in the household are still under the age of 18.

**SQL Code** In the following two sections of code we demonstrate the calculation of  $RFR_2$  or second-order re-finding rate. At this time, it is proposed that only second-order re-findings be used as measures under 1774. Future reports may include higher-order re-findings. Similar to  $RRR_i$ , we perform this calculation using our base working table of intakes (tbl\_intakes), (calendar\_dim), and a table of all children who appear to be associated with a given intake (tbl\_household\_children). For the sake of transparency, we again use four separate queries to complete this measurement:

1. A simple count of open intakes on the first day of 2010 with a finding indicated (RFR\_i\_start.sql),

```
declare @fystart int

set @fystart = 2010;

with hh_w_ch as
(
select
    ti.id_intake_fact
    , isnull(max(dateadd(yy, 18, thc.dt_birth))
        , dateadd(yy
            , 18
            , ti.inv_ass_start)) mx_dt_18
    , isnull(min(thc.dt_birth)
        , ti.inv_ass_start) mn_dt_br
from ca_ods.base.tbl_intakes ti
left join ca_ods.base.tbl_household_children thc
on ti.id_intake_fact = thc.id_intake_fact
and ti.id_case = thc.id_case
where
    cd_access_type in (1, 4)
    and (intake_rank = 1 or fl_founded_prior_any_legal = 0)
    and fl_founded_any_legal = 1
group by
    ti.id_intake_fact
```

```

,ti.inv_ass_start
)
select
  count(distinct id_intake_fact) pop_start
from hh_w_ch
where isnull(mx_dt_18,'3999-01-01') >=
  datefromparts(@fystart-1, 07, 01)
  and mn_dt_br < datefromparts(@fystart-1, 07, 01)

```

2. Two separate selections give us new first findings from 2010 to present (RFR\_i\_entry1.sql and RFR\_i\_entry2.sql), and

```

declare @fystart int
declare @fystop int

set @fystart = 2010
set @fystop = 2013;

select
  count(*) pop_enter
  ,month
from
  ca_ods.base.tbl_intakes ti
  left join ca_ods.dbo.calendar_dim cd
    on ti.inv_ass_start = calendar_date
where
  cd_access_type in (1, 4)
  and (intake_rank = 1 or fl_founded_prior_any_legal = 0)
  and fl_founded_any_legal = 1
  and inv_ass_start
    between datefromparts(@fystart-1, 07, 01)
    and datefromparts(@fystart, 06, 30)
group by
  month
order by
  month

```

```

declare @fystart int
declare @fystop int

set @fystart = 2010
set @fystop = 2013;

with ranked_subsequent_findings as
(
  select *
    ,rank() over
      (partition by id_case order by intake_rank desc) rank
  from ca_ods.base.tbl_intakes ti
  where
    cd_access_type in (1, 4)

```

```

and fl_founded_prior_any_legal = 1
and fl_founded_any_legal = 1
and inv_ass_start
between datefromparts(@fystart-1, 07, 01)
and datefromparts(@fystop, 06, 30)
)
select
count(*) pop_enter
,month
from ranked_subsequent_findings ti
left join ca_ods.dbo.calendar_dim cd
on ti.inv_ass_start = calendar_date
where
rank = 1
group by
month
order by
month

```

### 3. A selection of families who are “aging-out” of $P_{1PF<18}$ (RFR\_i\_exit.sql)

```

declare @fystart int
declare @fystop int

set @fystart = 2010
set @fystop = 2013;

with max_dt_18 as
(
select
ti.id_intake_fact
,isnull(max(dateadd(yy, 18, thc.dt_birth))
,dateadd(yy
,18
,ti.inv_ass_start)) mx_dt_18
from ca_ods.base.tbl_intakes ti
left join ca_ods.base.tbl_household_children thc
on ti.id_intake_fact = thc.id_intake_fact
and ti.id_case = thc.id_case
where
cd_access_type in (1, 4)
and (intake_rank = 1 or fl_founded_prior_any_legal = 0)
and fl_founded_any_legal = 1
group by
ti.id_intake_fact
,ti.inv_ass_start
having isnull(max(dateadd(yy, 18, thc.dt_birth))
,dateadd(yy
,18
,ti.inv_ass_start)) <=
datefromparts(@fystop, 06, 30)

```

```

)
select
  count(distinct id_intake_fact) pop_exit
  ,month
from ca_ods.dbo.calendar_dim cd
  left join max_dt_18 m18
    on mx_dt_18
      between dateadd(mm, -1, month) and month
where month between datefromparts(@fystart-1, 07, 01)
  and datefromparts(@fystop, 06, 30)
group by
  month
order by
  month

```

**R Code** We again load the four separate queries into R for further calculation:

```

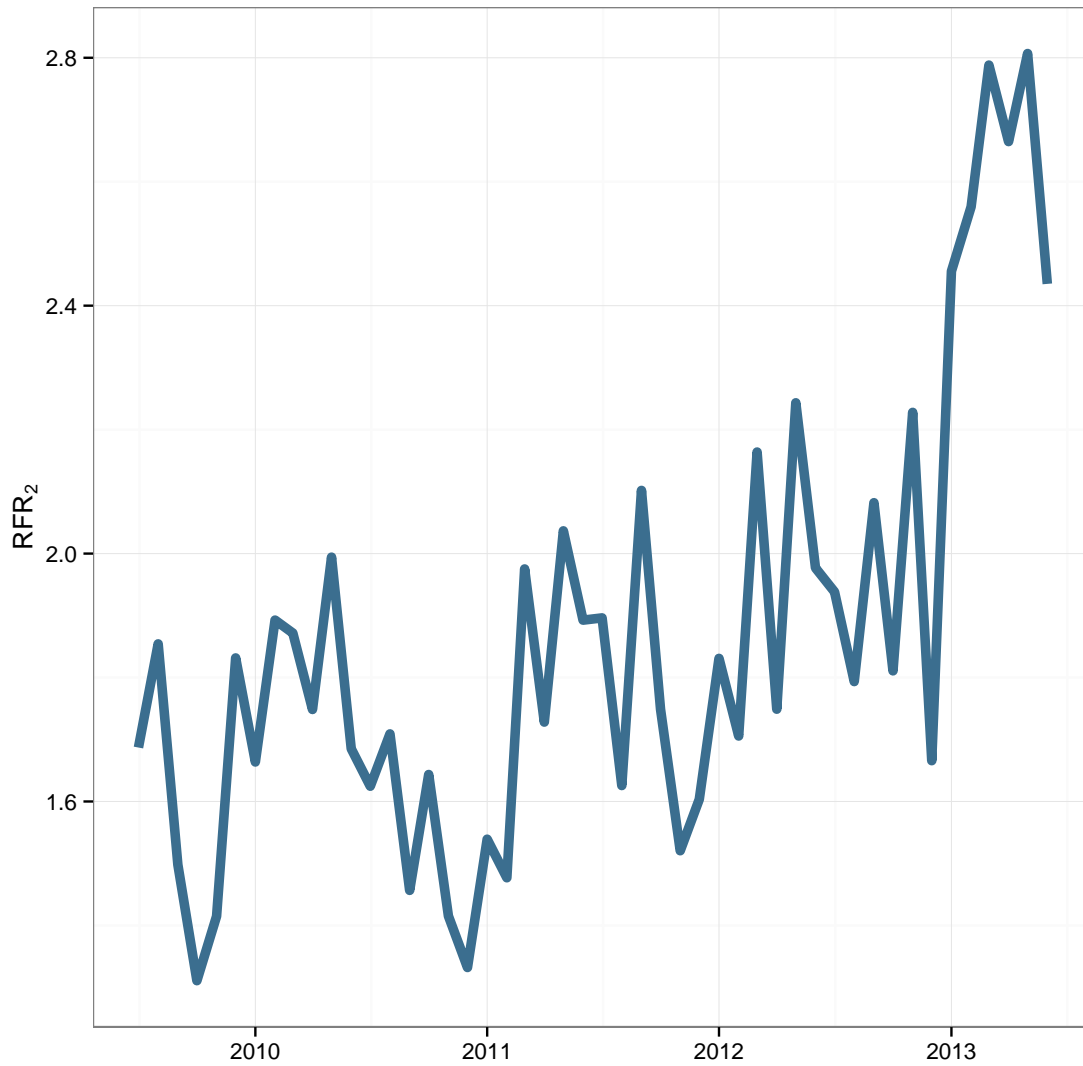
dat.RFR_i.s <- sqlQuery(cn, read.sql("RFR_i_start.sql"))
dat.RFR_i.e1 <- sqlQuery(cn, read.sql("RFR_i_entry1.sql"))
dat.RFR_i.e2 <- sqlQuery(cn, read.sql("RFR_i_entry2.sql"))
dat.RFR_i.ex <- sqlQuery(cn, read.sql("RFR_i_exit.sql"))

P_1PFlt18 <- dat.RFR_i.s$pop_start +
  dat.RFR_i.e1$pop_enter -
  dat.RFR_i.ex$pop_exit
I_F2 <- dat.RFR_i.e2$pop_enter
dat.RFR_i.e2$RFR2 <- (I_F2/P_1PFlt18)*1000

#plot time series
ggplot(dat.RFR_i.e2, aes(y=RFR2, x=month)) +
  geom_line(colour=poc_colors[1], size=2) +
  labs(y = expression(RFR[2]), x = "") +
  theme_bw()

```





**Foster-Care Finding Rate** The final safety measurement concerns the rate of findings per day of foster care. This is the first “care-day” measurement proposed in this document - measurements in which the denominator is a count of days of out-of-home care for some population as opposed to a count of some type of person.

**Definition** The Foster-Care Finding Rate ( $_{FC}FR$ ) follows the newly proposed CFSR measurements where  $_{FC}FR$  is defined as

$$_{FC}FR = \frac{_{FC}I_F}{c} \cdot 100,000 \quad (22)$$

where  $c$  represents the total number of care days provided by CA in a given year and  $_{FC}I_F$  represents the number of maltreatment findings for children in out-of-home care over that period. This measurement excludes children in foster care for less than 8 days and also excludes reports made within the first 7 days of the referral.

**SQL Code** In the following two sections of code we demonstrate the calculation of  $RRR_2$  or second-order re-referrals. For the sake of transparency and efficiency, we first load relevant data into a series of “global” temporary tables and then select from these tables. We plot a yearly time-series here as opposed to a monthly time-series for consistency with the CFSR measurements.

```
declare @fystart int
declare @fystop int

set @fystart = 2000
set @fystop = 2013;

if object_id('tempdb..##calendar_dim_sub') is not null
    drop table ##calendar_dim_sub

select distinct
    id_calendar_dim
    ,state_fiscal_yyyy
into ##calendar_dim_sub
from ca_ods.dbo.calendar_dim
where
    state_fiscal_yyyy between @fystart and @fystop;

if object_id('tempdb..##placement_prep') is not null
    drop table ##placement_prep

select distinct
    dbo.IntDate_to_CalDate(af.id_calendar_dim_incident)
        as incident_date
    ,af.id_calendar_dim_incident
    ,rp.removal_dt
    ,convert(int, convert(varchar, rp.removal_dt, 112))
        as removal_dt_int
    ,convert
    (
        int
    ,convert
        (varchar
        ,case
            when iif(rp.[18bday] < rp.discharge_dt
                ,rp.[18bday]
                ,rp.discharge_dt) = '9999-12-31'
            then datefromparts(@fystop, 06, 30)
            else iif(rp.[18bday] < rp.discharge_dt
                ,rp.[18bday]
                ,rp.discharge_dt)
            end
        ,112
    )
    ) discharg_frc_18_int
```

```

, rp.id_removal_episode_fact as care_days_cntr
, af.id_abuse_fact as substantiated_allegation_cntr
into ##placement_prep
from ca_ods.base.rptPlacement rp
left join ca_ods.dbo.abuse_fact af
on af.id_prsn_vctm = rp.child
and datediff(dd, removal_dt
, dbo.IntDate_to_CalDate(af.id_calendar_dim_incident)) > 7
and case
when iif(rp.[18bday] < rp.discharge_dt
, rp.[18bday], rp.discharge_dt) = '9999-12-31'
then datediff(dd, rp.removal_dt
, datefromparts(@fystop, 06, 30))
else datediff(dd, rp.removal_dt
, iif(rp.[18bday] < rp.discharge_dt
, rp.[18bday]
, rp.discharge_dt))
end + 1 > 7

if object_id('tempdb..##care_days') is not null
drop table ##care_days

select
cd.id_calendar_dim
, cd.state_fiscal_yyyy fiscal_yr
, count(rp1.care_days_cntr) care_days
into ##care_days
from ##calendar_dim_sub cd
left join ##placement_prep rp1
on cd.id_calendar_dim between removal_dt_int
and discharg_frc_18_int
where
cd.state_fiscal_yyyy between @fystart and @fystop
group by
cd.id_calendar_dim
, cd.state_fiscal_yyyy

if object_id('tempdb..##substantiations') is not null
drop table ##substantiations

select
cd.id_calendar_dim
, cd.state_fiscal_yyyy fiscal_yr
, count(rp2.substantiated_allegation_cntr)
as substantiated_allegations
into ##substantiations
from ##calendar_dim_sub cd
left join ##placement_prep rp2
on cd.id_calendar_dim = rp2.id_calendar_dim_incident

```

```

where
  cd.state_fiscal_yyyy between @fystart and @fystop
group by
  cd.id_calendar_dim
  ,cd.state_fiscal_yyyy

select
  sum(care_days) c
  ,sum(substantiated_allegations) FCI_F
  ,sum(substantiated_allegations)*100000.0/sum(care_days) FCFR
  ,cd.fiscal_yr
from ##care_days cd
  join ##substantiations s
    on cd.id_calendar_dim = s.id_calendar_dim
group by
  cd.fiscal_yr
order by
  cd.fiscal_yr

```

**R Code** In this chunk of code, we first execute the temporary table script (hence we need not assign the results of our first query to an object) and then select from our temporary tables with the select script. We then plot the data using ggplot2 as shown previously.

```

sqlQuery(cn, read.sql("FCFR_temp.sql"))

## [1] "42S22 207 [Microsoft][ODBC SQL Server Driver][SQL Server]Invalid
## [2] "42S22 207 [Microsoft][ODBC SQL Server Driver][SQL Server]Invalid
## [3] "42S22 207 [Microsoft][ODBC SQL Server Driver][SQL Server]Invalid
## [4] "[RODBC] ERROR: Could not SQLExecDirect 'declare @fystart int de

dat.FCFR <- sqlQuery(cn, read.sql("FCFR_select.sql"))

#plot time series
ggplot(dat.FCFR, aes(y=FCFR, x=fiscal_yr)) +
  geom_line(colour=poc_colors[1], size=2) +
  labs(y = expression(paste({}[FC], "FR")), x = "") +
  theme_bw()

## Error: ggplot2 doesn't know how to deal with data of class
character

```

### 3.4. Measurements Related to Permanency

#### 3.4.1. Placement stability for children placed in out-of-home care

Order-specific mobility rates according to length of stay

Order-transition-specific mobility rates according to length of stay

### 3.4.2. Length of time to permanence for children in out-of-home care

**Probability of Remaining in Care for 1 Year** This measurement is similar to the proposed CFSR measurement requiring exit to permanency within 12 months. However, instead of calculating a specific percentage, we propose that a Kaplan-Meier estimator be used to better account for right censoring of children who either age-out of the system or exit to a form of permanency other than adoption, reunification, or some form of guardianship.

**Definition** The Kaplan-Meier estimator for a federally recognized form of permanency  $P$  is given as

$${}_P\hat{S}(t) = \prod_{t_i < t} \frac{n_i - p_i}{n_i} \quad (23)$$

where  $n_i$  is the number still in care just prior to time  $t_i$ , and  $p_i$  is the number of permanency events at  $t_i$ . Thus, the probability of exit within 1 year is given as  ${}_P\hat{S}(365)$ . A reasonable approximation of the new CFSR measurement is given as

$$(1 - {}_P\hat{S}(365)) \cdot 100. \quad (24)$$

**SQL Code** POC is currently working on an implementation of a Kaplan-Meier estimator in SQL server. For our present purposes, however, we will select the relevant data with SQL code and calculate  ${}_P\hat{S}(365)$  using `survival` package in R. The relevant SQL code is as follows:

```
declare @fystart int
declare @fystop int
declare @last_obs datetime
set @fystart = 2000
set @fystop = 2013;

select
  child id_prsn_child
, case
  when rp.discharge_dt = '9999-12-31'
  then datediff(dd
    , rp.removal_dt
    , datefromparts(@fystop, 06, 30))
  when rp.[18bday] < rp.discharge_dt
  then datediff(dd
    , rp.removal_dt
    , rp.[18bday])
  else datediff(dd
    , rp.removal_dt
    , rp.discharge_dt)
end t
```

```

,case
  when rp.discharge_dt <> '9999-12-31'
    and rp.[18bday] > rp.discharge_dt
    and toe.discharge_type in ('Reunification'
                              , 'Adoption'
                              , 'Guardianship')

    then 1
    else 0
  end p
,cd1.state_fiscal_yyyy fiscal_yr
from ca_ods.base.rptPlacement rp
left join dbo.ref_lookup_cd_discharge_type_exits toe
  on toe.cd_discharge_type=rp.cd_discharge_type
join ca_ods.dbo.calendar_dim cd1
  on rp.removal_dt = cd1.calendar_date
where
  cd1.state_fiscal_yyyy between @fystart and @fystop
  and case
    when rp.discharge_dt = '9999-12-31'
    then datediff(dd
      ,rp.removal_dt
      ,datefromparts(@fystop, 06, 30))
    when rp.[18bday] < rp.discharge_dt
    then datediff(dd
      ,rp.removal_dt
      , rp.[18bday])
    else datediff(dd
      ,rp.removal_dt
      ,rp.discharge_dt)
    end + 1 > 7
order by t

```

**R Code** The following code allows us to load the data using the aforementioned SQL script as we have done previously. We then set our model using the R syntax for event history models. This model statement indicates our time variable  $t$  and our permanency variable  $p$ . We set the model “equal to” `fiscal_yr` which tells R that we want to calculate separate survival functions for each entry-year. The fitted survival functions are assigned to an object named `fit`. We then create a new dataframe called `dat` which we initialize with a `fiscal_yr` column indicating the fiscal year in question and a `P_hat_S365` column into which we will loop values of  $\hat{P}_S(365)$  for each fiscal year.

```

require(survival)

## Loading required package: survival
## Loading required package: splines

dat.P_hat_S <- sqlQuery(cn, read.sql("p_hat_S.sql"))

model <- Surv(t,p) ~ fiscal_yr

```

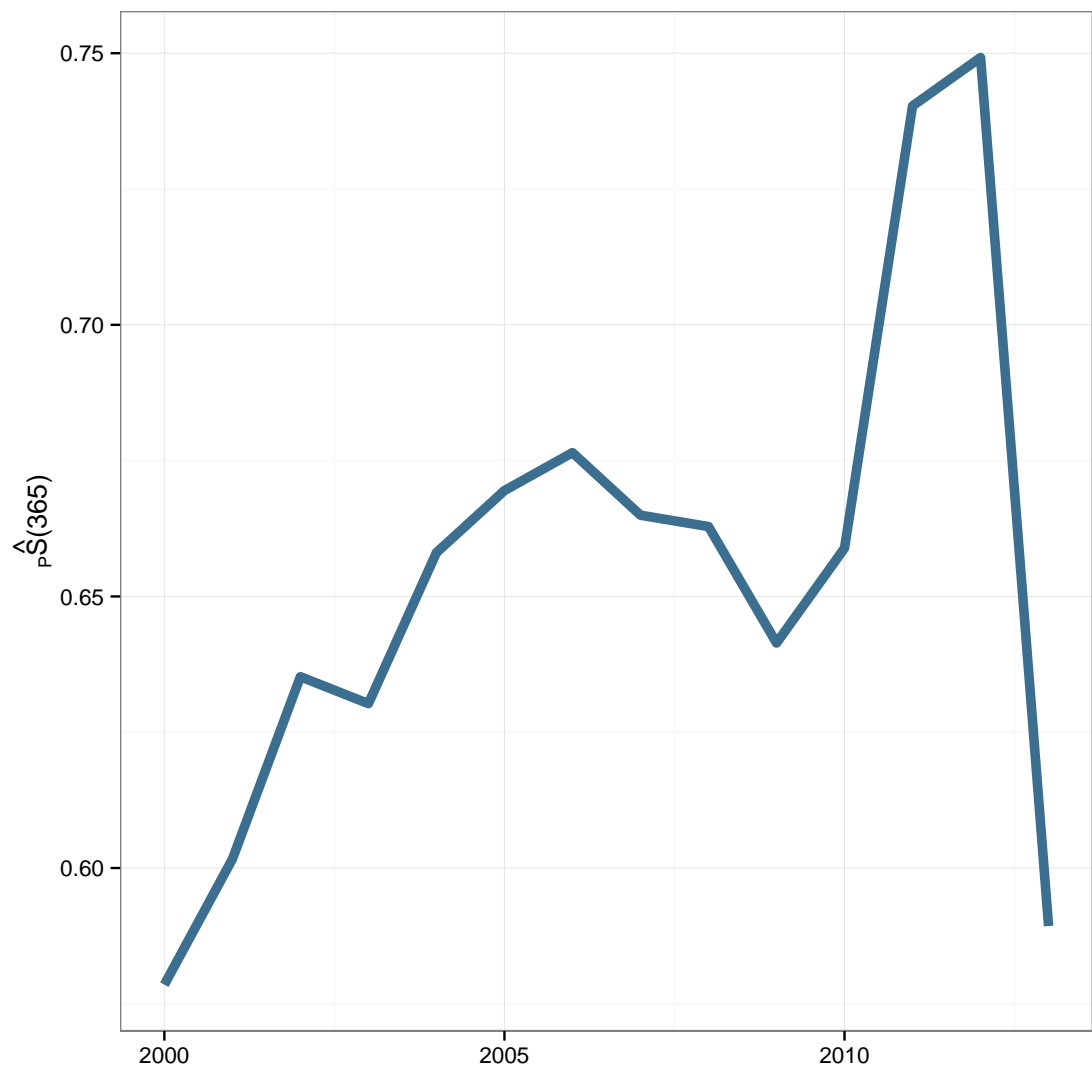
```

fit <- survfit(model, dat=dat.P_hat_S)

dat <- data.frame(fiscal_yr = 2000:2013
                  , P_hat_S365 = rep(NA, length(fit$strata)))
for(i in 1:length(fit$strata)){
  sum_i <- summary(fit[i])
  idx <- which.min(abs(summary(fit[i])$time-365))
  dat$P_hat_S365[i] <- sum_i$surv[idx]
}

#plot time series
ggplot(dat, aes(y=P_hat_S365, x=fiscal_yr)) +
  geom_line(colour=poc_colors[1], size=2) +
  labs(y = expression(paste({}[P]
                             , hat(S)
                             , "("
                             , 365
                             , ")"))
       , x = "") +
  theme_bw()

```



```
require(survival)
dat.A_hat_S <- sqlQuery(cn, read.sql("A_hat_S.sql"))

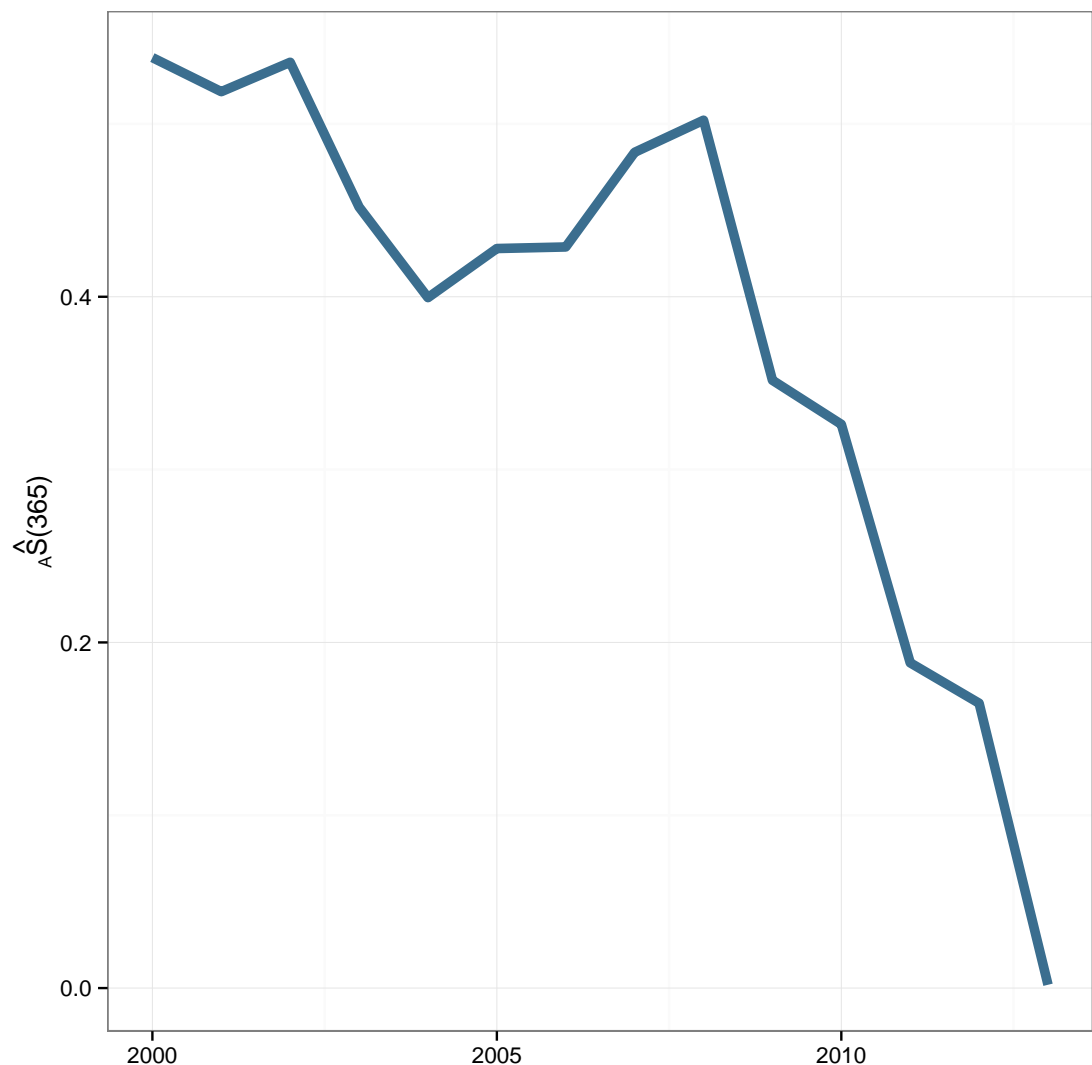
model <- Surv(t,a) ~ fiscal_yr
fit <- survfit(model, dat=dat.A_hat_S)

dat <- data.frame(fiscal_yr = 2000:2013
                  ,A_hat_S365 = rep(NA,length(fit$strata)))
for(i in 1:length(fit$strata)){
  sum_i <- summary(fit[i])
  idx <- which.min(abs(summary(fit[i])$time-365))
  dat$A_hat_S365[i] <- sum_i$surv[idx]
}

#plot time series
ggplot(dat, aes(y=A_hat_S365, x=fiscal_yr)) +
```



```
geom_line(colour=poc_colors[1], size=2) +
labs(y = expression(paste({}[A]
                        ,hat(S)
                        , " ("
                        , 365
                        , ") ") )
      , x = "") +
theme_bw()
```



### 3.4.3. Safe reunification of children placed in out-of-home care

#### Reentry rate according to permanency duration

**Probability of Permanency for 1 Year** This measurement is similar to the proposed CFSR measurement examining reentry within 12 months. However, instead of calculating

a specific percentage, we propose that a Kaplan-Meier estimator be used to better account for right censoring of children who either age-out of the system or exit to a form of permanency other than adoption, reunification, or some form of guardianship.

**Definition** The Kaplan-Meier estimator for reentry  $R$  from a federally recognized form of permanency is given as

$${}_R\hat{S}(t) = \prod_{t_i < t} \frac{n_i - r_i}{n_i} \quad (25)$$

where  $n_i$  is the number still in permanency just prior to time  $t_i$ , and  $r_i$  is the number of reentry events at  $t_i$ . Thus, the probability of exit within 1 year is given as  ${}_R\hat{S}(365)$ . A reasonable approximation of the new CFSR measurement is given as

$$(1 - {}_R\hat{S}(365)) \cdot 100. \quad (26)$$

**SQL Code** As mentioned, POC is currently working on an implementation of a Kaplan-Meier estimator in SQL server. For our present purposes, however, we will select the relevant data with SQL code and calculate  ${}_R\hat{S}(365)$  as we did for  ${}_P\hat{S}(365)$ . The relevant SQL code is as follows:

```
declare @fystart int
declare @fystop int
set @fystart = 2000
set @fystop = 2013;

select
  child id_prsn_child
, case
  when rp.[18bday] < next_reentry_date
  then datediff(dd, rp.discharge_dt, rp.[18bday]) + 1
  else datediff(dd, rp.discharge_dt, isnull(rp.next_reentry_date, datefrompart
end t
, case
  when rp.next_reentry_date is null and rp.[18bday] < datefromparts(@fystop,
  then 0
  else iif(rp.next_reentry_date is not null, 1, 0)
end r
, cd1.state_fiscal_yyyy fiscal_yr
from ca_ods.base.rptPlacement rp
left join dbo.ref_lookup_cd_discharge_type_exits toe
on toe.cd_discharge_type=rp.cd_discharge_type
join ca_ods.dbo.calendar_dim cd1
on convert(int
, convert(
  varchar
, case
```

```

        when rp.discharge_dt = '9999-12-31'
        then datefromparts(@fystop, 06, 30)
        when rp.[18bday] < rp.discharge_dt
        then rp.[18bday]
        else rp.discharge_dt
    end
    ,112
    )
    ) = cd1.id_calendar_dim
where
cd1.state_fiscal_yyyy between @fystart and @fystop
and rp.discharge_dt <> '9999-12-31'
and case
    when rp.[18bday] < next_reentry_date
    then datediff(dd, rp.discharge_dt, rp.[18bday]) + 1
    else datediff(dd, rp.discharge_dt, isnull(rp.next_reentry_date, datefromparts(@fystop, 06, 30)))
    end >= 0
and case
    when rp.discharge_dt <> '9999-12-31'
        and rp.[18bday] > rp.discharge_dt
        and toe.discharge_type in ('Reunification'
                                   , 'Adoption'
                                   , 'Guardianship')
    then 1
    else 0
    end = 1
and case
    when rp.discharge_dt = '9999-12-31'
    then datediff(dd
        ,rp.removal_dt
        ,datefromparts(@fystop, 06, 30))
    when rp.[18bday] < rp.discharge_dt
    then datediff(dd
        ,rp.removal_dt
        , rp.[18bday])
    else datediff(dd
        ,rp.removal_dt
        ,rp.discharge_dt)
    end + 1 > 7
order by t

```

**R Code** The following code allows us to load the data using the aforementioned SQL script as we have done previously. We then set our model using the R syntax for event history models. This model statement indicates our time variable `t` and our reentry variable `r`. Similar to the above, We set the model “equal to” `fiscal_yr` which tells R that we want to calculate separate survival functions for each entry-year. The fitted survival functions are assigned to an object named `fit`. We then create a new dataframe called `dat` which we initialize with a `fiscal_yr` column indicating the fiscal year in

question and a `R_hat_S365` column into which we will loop values of  $\hat{S}_R(365)$  for each fiscal year.

```
require(survival)
dat.R_hat_S <- sqlQuery(cn, read.sql("R_hat_S.sql"))

model <- Surv(t,r) ~ fiscal_yr
fit <- survfit(model, dat=dat.R_hat_S)

dat <- data.frame(fiscal_yr = 2000:2013
                  , R_hat_S365 = rep(NA, length(fit$strata)))
for(i in 1:length(fit$strata)){
  sum_i <- summary(fit[i])
  idx <- which.min(abs(summary(fit[i])$time-365))
  dat$R_hat_S365[i] <- sum_i$surv[idx]
}

#plot time series
ggplot(dat, aes(y=R_hat_S365, x=fiscal_yr)) +
  geom_line(colour=poc_colors[1], size=2) +
  labs(y = expression(paste({}[R]
                             , hat(S)
                             , "("
                             , 365
                             , ")"))
       , x = "") +
  theme_bw()
```

