GVisualization Planning for Heracles

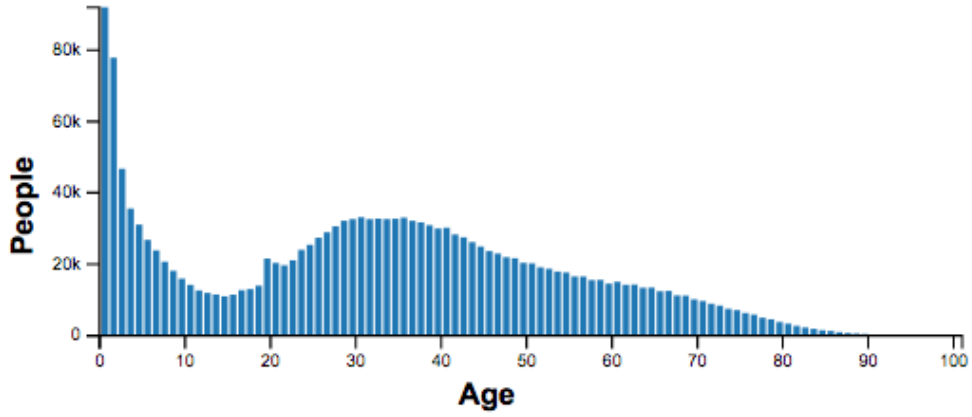
- Age
    - Number of persons by age, with age at cohort start
        - Histogram
        - 1800
    - ~~Distribution of age at cohort start~~
        - ~~Box plot underneath histogram~~
        - ~~1801~~
- Age by Race
- ~~Dist of Age at cohort start by Gender~~
    - ~~1802~~
    - ~~Box Plot~~

- Ethnicity
- ~~Number of persons by duration from cohort start to cohort end with index in middle~~
    - ~~Line chart with bar in middle~~
    - ~~1805 (left side) + 1806 (right side)~~
- ~~Number of persons by cohort start by gender by age~~
    - ~~1814~~
    - ~~Trellis plot - calculate prevalence based on Achilles data~~
- ~~Number of persons by cohort start month~~
    - ~~1815~~
    - ~~Line plot~~
- ~~Distribution of age at cohort start by cohort start year~~
    - ~~1803~~
    - ~~Box plot~~
- Treemap of all conditions
    - 1820/1821
- Treemap of procedures (similar to 1820)
    - 1830/1831
- Number of persons by duration from cohort start to first occurrence of drug era
    - 1870
- Number of events by duration from cohort start to all occurrences of drug era
    - 1871
- ~~Number of persons from cohort start to end~~
    - ~~1804 - line graph~~

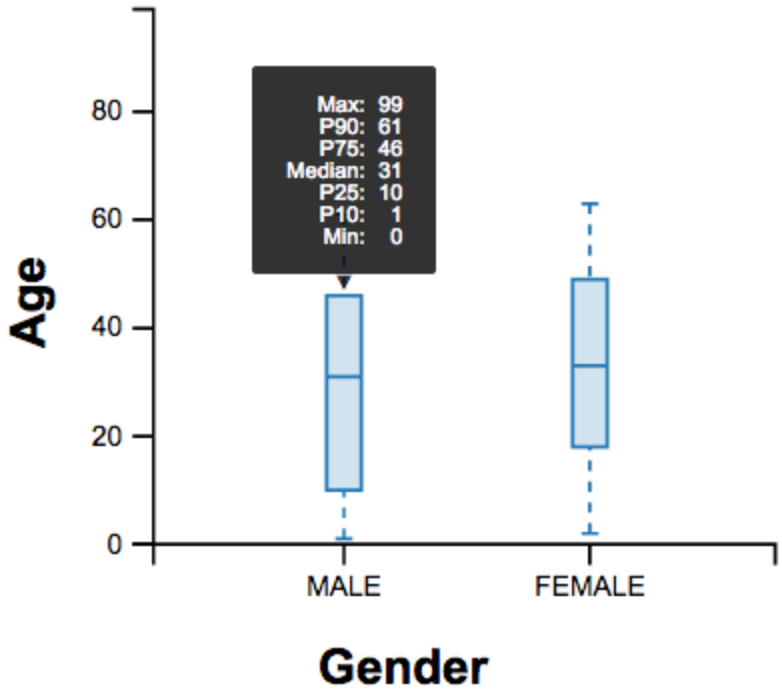
1807 - plain line chart
1800 - bar o

Legend:

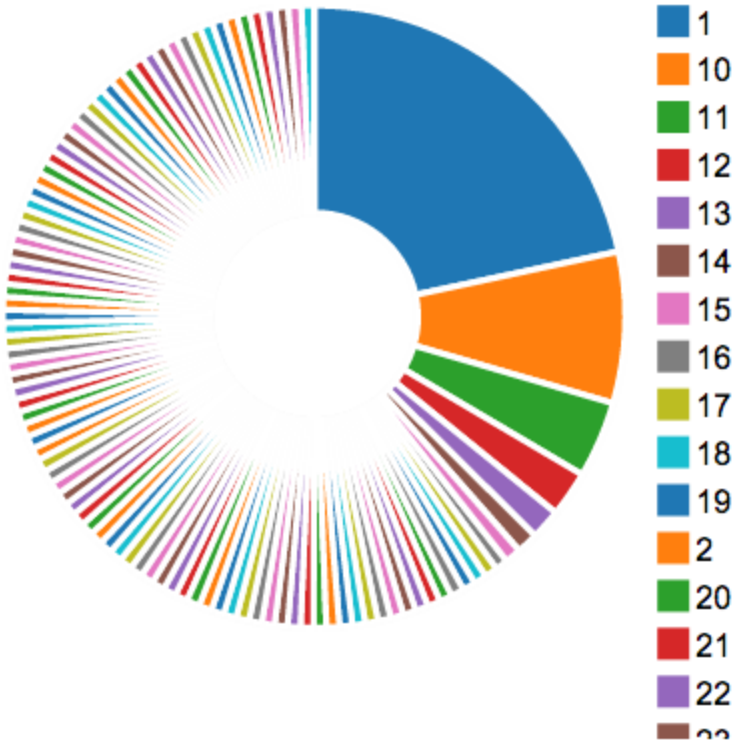
Histogram



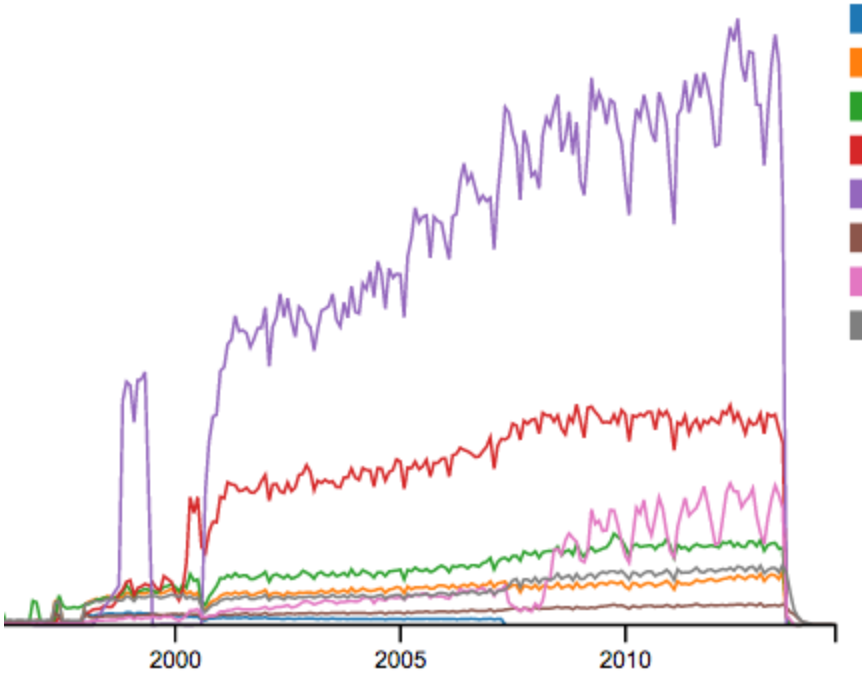
Box Plot (can also do horizontal box plots to pull in the detailed data)



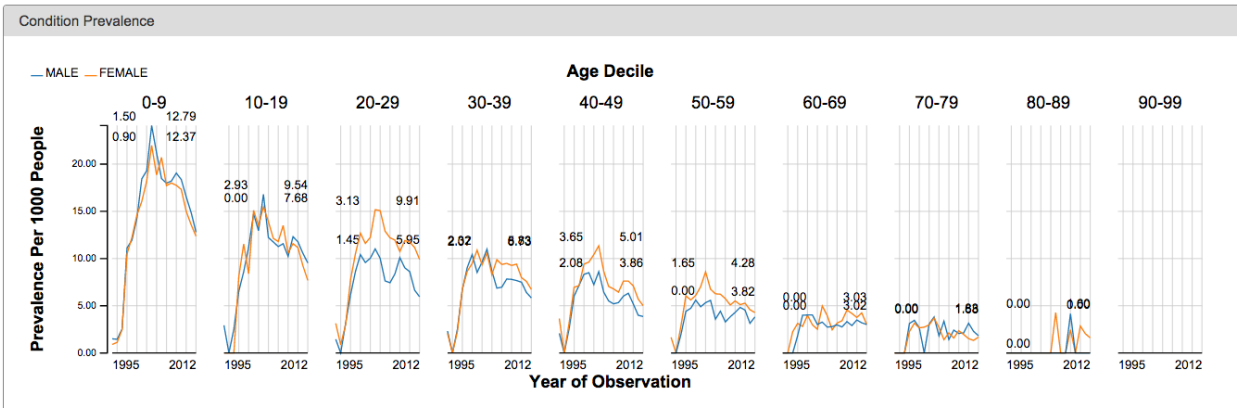
Donut Plot



Line Plot



Trellis Plot



1800 Number of persons by age, with age at cohort start (histogram)

1801 Distribution of age at cohort start (histogram vs box plot) - horizontal box plot under the histogram plot

1802 Distribution of age at cohort start by gender (box plot – same as age by gender) - NO QUERY

1803 Distribution of age at cohort start by cohort start year (box plot with categories as years) - NO QUERY

1804 Number of persons by duration from cohort start to cohort end, in 30d increments (line graph-persons with cont obs by month - akin to observation length) - NO QUERY

1805 Number of persons by duration from observation start to cohort start, in 30d increments (line graph vs bar - with index date in middle; t-index, t+index)

1806 Number of persons by duration from cohort start to observation end, in 30d increments (right half of 1805)

1807 Number of persons by duration from cohort end to observation end, in 30d increments (line chart, index date in middle) - ? include, would need to have multiple axes in order to accomodate - ?usability - NO QUERY

1808 Distribution of duration (days) from cohort start to cohort end (box plot with x-axis cohort start to cohort end) - NO QUERY

1809 Distribution of duration (days) from cohort start to cohort end, by gender (box plot) - NO QUERY

1810 Distribution of duration (days) from cohort start to cohort end, by age decile (box plot) - NO QUERY

1811 Distribution of duration (days) from observation start to cohort start (a la 1808) - NO QUERY

1812 Distribution of duration (days) from cohort start to observation end (a la 1808) - NO QUERY

1813 Distribution of duration (days) from cohort end to observation end (a la 1808) - NO QUERY

1814 Number of persons by cohort start year by gender by age decile (trellis plot - calculate prevalence by pulling from achilles data)

1815 Number of persons by cohort start month (line plot - like COndition prevalence per month - could use denominator from Achilles)

1816 Number of persons by number of cohort periods (donut plot, like observation periods per person but cohort periods per person, may need to group to get representation) ? if # categories gets high, graph may not look good but it's a good data - follow up re data quality issue - NO QUERY

1820 Number of persons by duration from cohort start to first occurrence of condition occurrence, by condition_concept_id (Treemap with size based on prevalence within the cohort and color with ratio of before / after (precursors are darker while consequences are lighter) → details to line with index - (can the treemap have dynamic coloring based on the range of observed values, so for example you can look at cohort vs overall database - would need to shift color)

1821 Number of events by duration from cohort start to all occurrences of condition occurrence, by condition_concept_id (a la 1820 with details to line with index)

1830 Number of persons by duration from cohort start to first occurrence of procedure occurrence, by procedure_concept_id (a la 1820 with details to line with index)

1831 Number of events by duration from cohort start to all occurrences of procedure occurrence, by procedure_concept_id (Treemap with size based on prevalence within the cohort and color with ratio of before / after (precursors are darker while consequences are lighter) → details to line with index

Number of persons by duration from cohort start to first occurrence of drug exposure, by drug_concept_id (Treemap with size based on prevalence within the cohort and color with ratio of before / after (precursors are darker while
1840 consequences are lighter) → details to line with index - NO QUERY

Number of events by duration from cohort start to all occurrences of drug exposure,
1841 by drug_concept_id (a la 1840 → details to line with index) - NO QUERY

Number of persons by duration from cohort start to first occurrence of observation, by observation_concept_id Treemap with size based on prevalence within the cohort and color with ratio of before / after (precursors are darker while consequences are
1850 lighter) → details to line with index - NO QUERY

Number of events by duration from cohort start to all occurrences of observation, by
1851 observation_concept_id - NO QUERY

Number of persons by duration from cohort start to first occurrence of condition era,
1860 by condition_concept_id - NO QUERY

Number of events by duration from cohort start to all occurrences of condition era,
1861 by condition_concept_id - NO QUERY

Number of persons by duration from cohort start to first occurrence of drug era, by
1870 drug_concept_id

Number of events by duration from cohort start to all occurrences of drug era, by
1871 drug_concept_id

Export SQL

```
/*****************
HERACLES sql script for JSON export

Patrick Ryan

last updated: 25 Jan 2015


******************/

/*
Parameter used:

{DEFAULT @cdmSchema = 'cdmSchema'}  --cdmSchema = @cdmSchema
{DEFAULT @resultsSchema = 'resultsSchema'}   --resultsSchema = @resultsSchema
{DEFAULT @heraclesResultsTable = 'heracles_results'}  --heraclesResultsTable =
@heraclesResultsTable
{DEFAULT @heraclesResultsDistTable = 'heracles_results_dist'}
--heraclesResultsDistTable = @heraclesResultsDistTable
{DEFAULT @achillesResultsTable = 'achilles_results'}  --achillesResultsTable =
@achillesResultsTable
{DEFAULT @achillesResultsDistTable = 'achilles_results_dist'}
--achillesResultsDistTable = @achillesResultsDistTable
{DEFAULT @cohortDefinitionId = 1  --cohortDefinitionId = @cohortDefinitionId
{DEFAULT @minCovariatePersonCount = 500 --minCovariatePersonCount =
@minCovariatePersonCount
{DEFAULT @minIntervalPersonCount = 1000 --@IntervalPersonCount =
@minIntervalPersonCount
*/



/*

Cohort characteristics
--age distribution
--gender
--observation period
*/


--age at index
select cast(stratum_1 as integer) as age_at_index,
       count_value as num_persons
from @resultsSchema.dbo.@heraclesResultsTable
```

```
where analysis_id in (1800)
and cohort_definition_id in (@cohortDefinitionId)


--age at index distribution
select count_value, min_value, max_value, avg_value, stdev_value, p10_value,
p25_value, median_value, p75_value, p90_value
from @resultsSchema.dbo.@heraclesResultsDistTable
where analysis_id in (1801)
and cohort_definition_id in (@cohortDefinitionId)


--gender
select c1.concept_id,
       c1.concept_name,
       hr1.count_value as num_persons
from
(
select cast(stratum_1 as integer) as concept_id,
       count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (2)
and cohort_definition_id in (@cohortDefinitionId)
) hr1
inner join
@cdmSchema.dbo.concept c1
on hr1.concept_id = c1.concept_id
;


--race
select c1.concept_id,
       c1.concept_name,
       hr1.count_value as num_persons
from
(
select cast(stratum_1 as integer) as concept_id,
       count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (4)
and cohort_definition_id in (@cohortDefinitionId)
) hr1
inner join
@cdmSchema.dbo.concept c1
on hr1.concept_id = c1.concept_id
;
```

```
--ethnicity
select c1.concept_id,
       c1.concept_name,
       hr1.count_value as num_persons
from
(
select cast(stratum_1 as integer) as concept_id,
       count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (5)
and cohort_definition_id in (@cohortDefinitionId)
) hr1
inner join
@cdmSchema.dbo.concept c1
on hr1.concept_id = c1.concept_id
;


--observation period time, relative to index
select hr1.cohort_definition_id,
             hr1.duration,
             hr1.count_value,
             1.0*hr1.count_value/t1.count_value as pct_persons
from
(
select cohort_definition_id,
             -1* cast(stratum_1 as integer)*30 as duration,
             sum(count_value) over (partition by cohort_definition_id order by -1*
cast(stratum_1 as integer)*30 asc) as count_value
       from
       @resultsSchema.dbo.@heraclesResultsTable
       where analysis_id in (1805)
       and cohort_definition_id in (@cohortDefinitionId)
       and cast(stratum_1 as integer) > 0

       union

       select hr1.cohort_definition_id,
              cast(hr1.stratum_1 as integer)*30 as duration,
              t1.count_value - sum(hr1.count_value) over (partition by
hr1.cohort_definition_id order by cast(hr1.stratum_1 as integer)*30 asc) as
count_value
       from
       @resultsSchema.dbo.@heraclesResultsTable hr1
       inner join
       (select cohort_definition_id, sum(count_value) as count_value
       from @resultsSchema.dbo.@heraclesResultsTable
       where analysis_id = 1806
```

```
	and cohort_definition_id in (@cohortDefinitionId)
	group by cohort_definition_id) t1
	on hr1.cohort_definition_id = t1.cohort_definition_id
	where hr1.analysis_id in (1806)
	and hr1.cohort_definition_id in (@cohortDefinitionId)
) hr1,
(select count_value from @resultsSchema.dbo.@heraclesResultsTable where analysis_id =
1) t1
;



--prevalence by month

select hr1.stratum_1 as calendar_month,
	substring(hr1.stratum_1,1,4) as calendar_year,
	substring(hr1.stratum_1,5,2) as calendar_month_index,
	hr1.count_value as num_persons,
	round(1000*(1.0*hr1.count_value / t1.count_value),5) as y_prevelance_1000pp
from (select stratum_1, count_value
	from @resultsSchema.dbo.@heraclesResultsTable
	where analysis_id in (1815)
	and cohort_definition_id in (@cohortDefinitionId)
) hr1
	inner join
(
	select stratum_1, count_value from @resultsSchema.dbo.@achillesResultsTable
where analysis_id = 117
) t1
on hr1.stratum_1 = t1.stratum_1



--prevalence by year, gender, sex

select hr1.cohort_definition_id,
	hr1.index_year,
	c1.concept_name as gender_concept_name,
	cast(hr1.age_decile*10 as varchar) + '-' + cast((hr1.age_decile+1)*10-1 as
varchar) as age_decile,
	hr1.count_value as num_persons,
	round(1000*(1.0*hr1.count_value / t1.count_value),5) as y_prevelance_1000pp
from (select cohort_definition_id,
	cast(stratum_1 as integer) as index_year,
	cast(stratum_2 as integer) as gender_concept_id,
	cast(stratum_3 as integer) as age_decile,
	count_value
	from @resultsSchema.dbo.@heraclesResultsTable
```

```
	where analysis_id in (1814)
	and cohort_definition_id in (@cohortDefinitionId)
	and stratum_2 in (8507,8532)
	and stratum_3 >= 0 and stratum_4 <10
) hr1
	inner join
(
	select cast(stratum_1 as integer) as index_year,
	cast(stratum_2 as integer) as gender_concept_id,
	cast(stratum_3 as integer) as age_decile,
	count_value
	from @resultsSchema.dbo.@achillesResultsTable
	where analysis_id = 116
) t1
on hr1.index_year = t1.index_year
and hr1.gender_concept_id = t1.gender_concept_id
and hr1.age_decile = t1.age_decile
inner join
@cdmSchema.dbo.concept c1
on hr1.gender_concept_id = c1.concept_id







/*
CONDITION_OCCURRENCE

--treemap of all conditions
--analysis_id: 1820
--size - prevalence of condition
--color:  risk difference of prevalence before / after index
*/


select   concept_hierarchy.concept_id,
	isNull(concept_hierarchy.soc_concept_name,'NA') as soc_concept_name,
	isNull(concept_hierarchy.hlgt_concept_name,'NA') as hlgt_concept_name,
	isNull(concept_hierarchy.hlt_concept_name,'NA') as hlt_concept_name,
	isNull(concept_hierarchy.pt_concept_name,'NA') as pt_concept_name,
	isNull(concept_hierarchy.snomed_concept_name,'NA') as concept_name,
  isNull(concept_hierarchy.soc_concept_name,'NA') + '||' +
isNull(concept_hierarchy.hlgt_concept_name,'NA') + '||' +
isNull(concept_hierarchy.hlt_concept_name,'NA') + '||' +
isNull(concept_hierarchy.pt_concept_name,'NA') + '||' +
isNull(concept_hierarchy.snomed_concept_name,'NA') as concept_path,
```

```
	1.0*hr1.num_persons / denom.count_value as percent_persons,
	1.0*hr1.num_persons_before / denom.count_value as percent_persons_before,
	1.0*hr1.num_persons_after / denom.count_value as percent_persons_after,
	1.0*(hr1.num_persons_after - hr1.num_persons_before)/denom.count_value as
risk_diff_after_before,
	log(1.0*(hr1.num_persons_after + 0.5) / (hr1.num_persons_before + 0.5)) as
logRR_after_before
from
(select cast(stratum_1 as integer) as concept_id,
	sum(count_value) as num_persons,
	sum(case when stratum_2 < 0 then count_value else 0 end) as
num_persons_before,
	sum(case when stratum_2 > 0 then count_value else 0 end) as num_persons_after
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1820) --first occurrence of condition
and cohort_definition_id in (@cohortDefinitionId)
group by cast(stratum_1 as int)
) hr1
inner join
	(
		select snomed.concept_id,
			snomed.concept_name as snomed_concept_name,
			pt_to_hlt.pt_concept_name,
			hlt_to_hlgt.hlt_concept_name,
			hlgt_to_soc.hlgt_concept_name,
			soc.concept_name as soc_concept_name
		from
		(
		select concept_id, concept_name
		from @cdmSchema.dbo.concept
		where vocabulary_id = 'SNOMED'
		) snomed
		left join
			(select c1.concept_id as snomed_concept_id, max(c2.concept_id) as
pt_concept_id
			from
			@cdmSchema.dbo.concept c1
			inner join
			@cdmSchema.dbo.concept_ancestor ca1
			on c1.concept_id = ca1.descendant_concept_id
			and c1.vocabulary_id = 'SNOMED'
			and ca1.min_levels_of_separation = 1
			inner join
			@cdmSchema.dbo.concept c2
			on ca1.ancestor_concept_id = c2.concept_id
			and c2.vocabulary_id = 'MedDRA'
			group by c1.concept_id
			) snomed_to_pt
```

```
		on snomed.concept_id = snomed_to_pt.snomed_concept_id

		left join
			(select c1.concept_id as pt_concept_id, c1.concept_name as
pt_concept_name, max(c2.concept_id) as hlt_concept_id
			from
			@cdmSchema.dbo.concept c1
			inner join
			@cdmSchema.dbo.concept_ancestor ca1
			on c1.concept_id = ca1.descendant_concept_id
			and c1.vocabulary_id = 'MedDRA'
			and ca1.min_levels_of_separation = 1
			inner join
		  @cdmSchema.dbo.concept c2
			on ca1.ancestor_concept_id = c2.concept_id
			and c2.vocabulary_id = 'MedDRA'
			group by c1.concept_id, c1.concept_name
			) pt_to_hlt
		on snomed_to_pt.pt_concept_id = pt_to_hlt.pt_concept_id

		left join
			(select c1.concept_id as hlt_concept_id, c1.concept_name as
hlt_concept_name, max(c2.concept_id) as hlgt_concept_id
			from
			@cdmSchema.dbo.concept c1
			inner join
			@cdmSchema.dbo.concept_ancestor ca1
			on c1.concept_id = ca1.descendant_concept_id
			and c1.vocabulary_id = 'MedDRA'
			and ca1.min_levels_of_separation = 1
			inner join
			@cdmSchema.dbo.concept c2
			on ca1.ancestor_concept_id = c2.concept_id
			and c2.vocabulary_id = 'MedDRA'
			group by c1.concept_id, c1.concept_name
			) hlt_to_hlgt
		on pt_to_hlt.hlt_concept_id = hlt_to_hlgt.hlt_concept_id

		left join
			(select c1.concept_id as hlgt_concept_id, c1.concept_name as
hlgt_concept_name, max(c2.concept_id) as soc_concept_id
			from
			@cdmSchema.dbo.concept c1
			inner join
			@cdmSchema.dbo.concept_ancestor ca1
			on c1.concept_id = ca1.descendant_concept_id
			and c1.vocabulary_id = 'MedDRA'
			and ca1.min_levels_of_separation = 1
```

```
                    inner join
                    @cdmSchema.dbo.concept c2
                    on ca1.ancestor_concept_id = c2.concept_id
                    and c2.vocabulary_id = 'MedDRA'
                    group by c1.concept_id, c1.concept_name
                    ) hlgt_to_soc
                on hlt_to_hlgt.hlgt_concept_id = hlgt_to_soc.hlgt_concept_id

                left join @cdmSchema.dbo.concept soc
                 on hlgt_to_soc.soc_concept_id = soc.concept_id

        ) concept_hierarchy
        on hr1.concept_id = concept_hierarchy.concept_id
,
(select count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id = 1
) denom
;




/*
CONDITION_OCCURRENCE

--drilldown of when first condition occurs relative to index
--graph:  scatterplot
--analysis_id: 1820
--x:  time (30-day increments)
--y:  %
*/


select hr1.cohort_definition_id,
        'First' as record_type,
        c1.concept_id,
        c1.concept_name,
        hr1.duration,
        hr1.count_value,
        case when t1.count_value > 0 then 1.0*hr1.count_value / t1.count_value else 0
end as pct_persons
from
(select cohort_definition_id,
        cast(stratum_1 as integer) as concept_id,
        cast(stratum_2 as integer)*30 as duration,
        count_value
from @resultsSchema.dbo.@heraclesResultsTable
```

```
where analysis_id in (1820)
and cohort_definition_id in (@cohortDefinitionId)
) hr1
inner join
(
	select cohort_definition_id,
		-1* cast(stratum_1 as integer)*30 as duration,
		sum(count_value) over (partition by cohort_definition_id order by -1*
cast(stratum_1 as integer)*30 asc) as count_value
	from
	@resultsSchema.dbo.@heraclesResultsTable
	where analysis_id in (1805)
	and cohort_definition_id in (@cohortDefinitionId)
	and cast(stratum_1 as integer) > 0

	union

	select hr1.cohort_definition_id,
		cast(hr1.stratum_1 as integer)*30 as duration,
		t1.count_value - sum(hr1.count_value) over (partition by
hr1.cohort_definition_id order by cast(hr1.stratum_1 as integer)*30 asc) as
count_value
	from
	@resultsSchema.dbo.@heraclesResultsTable hr1
	inner join
	(select cohort_definition_id, sum(count_value) as count_value
	from @resultsSchema.dbo.@heraclesResultsTable
	where analysis_id = 1806
	and cohort_definition_id in (@cohortDefinitionId)
	group by cohort_definition_id) t1
	on hr1.cohort_definition_id = t1.cohort_definition_id
	where hr1.analysis_id in (1806)
	and hr1.cohort_definition_id in (@cohortDefinitionId)

) t1
on hr1.cohort_definition_id = t1.cohort_definition_id
and hr1.duration = t1.duration
inner join
(select cohort_definition_id,
	cast(stratum_1 as integer) as concept_id,
	sum(count_value) as count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1820)
and cohort_definition_id in (@cohortDefinitionId)
group by cohort_definition_id,
	cast(stratum_1 as integer)
having sum(count_value) > @minCovariatePersonCount
) ct1
```

```
on hr1.cohort_definition_id = ct1.cohort_definition_id
and hr1.concept_id = ct1.concept_id
inner join
@cdmSchema.dbo.concept c1
on hr1.concept_id = c1.concept_id
where t1.count_value > @minIntervalPersonCount

union


select hr1.cohort_definition_id,
	'All' as record_type,
	c1.concept_id,
	c1.concept_name,
	hr1.duration,
	hr1.count_value,
	case when t1.count_value > 0 then 1.0*hr1.count_value / t1.count_value else 0
end as pct_persons
from
(select cohort_definition_id,
	cast(stratum_1 as integer) as concept_id,
	cast(stratum_2 as integer)*30 as duration,
	count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1821)
and cohort_definition_id in (@cohortDefinitionId)
) hr1
inner join
(
	select cohort_definition_id,
		-1* cast(stratum_1 as integer)*30 as duration,
		sum(count_value) over (partition by cohort_definition_id order by -1*
cast(stratum_1 as integer)*30 asc) as count_value
	from
	@resultsSchema.dbo.@heraclesResultsTable
	where analysis_id in (1805)
	and cohort_definition_id in (@cohortDefinitionId)
	and cast(stratum_1 as integer) > 0

	union

	select hr1.cohort_definition_id,
		cast(hr1.stratum_1 as integer)*30 as duration,
		t1.count_value - sum(hr1.count_value) over (partition by
hr1.cohort_definition_id order by cast(hr1.stratum_1 as integer)*30 asc) as
count_value
	from
	@resultsSchema.dbo.@heraclesResultsTable hr1
```

```
	inner join
	(select cohort_definition_id, sum(count_value) as count_value
	from @resultsSchema.dbo.@heraclesResultsTable
	where analysis_id = 1806
	and cohort_definition_id in (@cohortDefinitionId)
	group by cohort_definition_id) t1
	on hr1.cohort_definition_id = t1.cohort_definition_id
	where hr1.analysis_id in (1806)
	and hr1.cohort_definition_id in (@cohortDefinitionId)

) t1
on hr1.cohort_definition_id = t1.cohort_definition_id
and hr1.duration = t1.duration
inner join
(select cohort_definition_id,
	cast(stratum_1 as integer) as concept_id,
	sum(count_value) as count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1821)
and cohort_definition_id in (@cohortDefinitionId)
group by cohort_definition_id,
	cast(stratum_1 as integer)
having sum(count_value) > @minCovariatePersonCount
) ct1
on hr1.cohort_definition_id = ct1.cohort_definition_id
and hr1.concept_id = ct1.concept_id
inner join
@cdmSchema.dbo.concept c1
on hr1.concept_id = c1.concept_id
where t1.count_value > @minIntervalPersonCount
;









/*
DRUG_ERA

--treemap of all conditions
--analysis_id: 1870
--size - prevalence of drug
--color:  risk difference of prevalence before / after index
*/
```

```
select   concept_hierarchy.concept_id,
       isnull(concept_hierarchy.rxnorm_ingredient_concept_name,'NA') as
ingredient_concept_name,
       isnull(concept_hierarchy.atc5_concept_name,'NA') as atc5_concept_name,
       isnull(concept_hierarchy.atc3_concept_name,'NA') as atc3_concept_name,
       isnull(concept_hierarchy.atc1_concept_name,'NA') as atc1_concept_name,
 isnull(concept_hierarchy.atc1_concept_name,'NA') + '||' +
       isnull(concept_hierarchy.atc3_concept_name,'NA') + '||' +
       isnull(concept_hierarchy.atc5_concept_name,'NA') + '||' +
       isnull(concept_hierarchy.rxnorm_ingredient_concept_name,'||') as concept_path,
       1.0*hr1.num_persons / denom.count_value as percent_persons,
       1.0*hr1.num_persons_before / denom.count_value as percent_persons_before,
       1.0*hr1.num_persons_after / denom.count_value as percent_persons_after,
       1.0*(hr1.num_persons_after - hr1.num_persons_before)/denom.count_value as
risk_diff_after_before,
       log(1.0*(hr1.num_persons_after + 0.5) / (hr1.num_persons_before + 0.5)) as
logRR_after_before
from
(select cast(stratum_1 as integer) as concept_id,
       sum(count_value) as num_persons,
       sum(case when stratum_2 < 0 then count_value else 0 end) as
num_persons_before,
       sum(case when stratum_2 > 0 then count_value else 0 end) as num_persons_after
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1870) --first occurrence of drug
and cohort_definition_id in (@cohortDefinitionId)
group by cast(stratum_1 as int)
) hr1
inner join
       (
       select rxnorm.rxnorm_ingredient_concept_id as concept_id,
                     rxnorm.rxnorm_ingredient_concept_name,
                     atc5_to_atc3.atc5_concept_name,
                     atc3_to_atc1.atc3_concept_name,
                     atc1.concept_name as atc1_concept_name
              from
              (
              select c2.concept_id as rxnorm_ingredient_concept_id,
                     c2.concept_name as RxNorm_ingredient_concept_name
              from
                     @cdmSchema.dbo.concept c2
                     where
                     c2.vocabulary_id = 'RxNorm'
                     and c2.concept_class_id = 'Ingredient'
              ) rxnorm
              left join
```

```
                    (select c1.concept_id as rxnorm_ingredient_concept_id,
max(c2.concept_id) as atc5_concept_id
                    from
                    @cdmSchema.dbo.concept c1
                    inner join
                    @cdmSchema.dbo.concept_ancestor ca1
                    on c1.concept_id = ca1.descendant_concept_id
                    and c1.vocabulary_id = 'RxNorm'
                    and c1.concept_class_id = 'Ingredient'
                    inner join
                    @cdmSchema.dbo.concept c2
                    on ca1.ancestor_concept_id = c2.concept_id
                    and c2.vocabulary_id = 'ATC'
                    and c2.concept_class_id = 'ATC 4th'
                    group by c1.concept_id
                    ) rxnorm_to_atc5
             on rxnorm.rxnorm_ingredient_concept_id =
rxnorm_to_atc5.rxnorm_ingredient_concept_id

             left join
                    (select c1.concept_id as atc5_concept_id, c1.concept_name as
atc5_concept_name, max(c2.concept_id) as atc3_concept_id
                    from
                    @cdmSchema.dbo.concept c1
                    inner join
                    @cdmSchema.dbo.concept_ancestor ca1
                    on c1.concept_id = ca1.descendant_concept_id
                    and c1.vocabulary_id = 'ATC'
                    and c1.concept_class_id = 'ATC 4th'
                    inner join
                    @cdmSchema.dbo.concept c2
                    on ca1.ancestor_concept_id = c2.concept_id
                    and c2.vocabulary_id = 'ATC'
                    and c2.concept_class_id = 'ATC 2nd'
                    group by c1.concept_id, c1.concept_name
                    ) atc5_to_atc3
             on rxnorm_to_atc5.atc5_concept_id = atc5_to_atc3.atc5_concept_id

             left join
                    (select c1.concept_id as atc3_concept_id, c1.concept_name as
atc3_concept_name, max(c2.concept_id) as atc1_concept_id
                    from
                    @cdmSchema.dbo.concept c1
                    inner join
                    @cdmSchema.dbo.concept_ancestor ca1
                    on c1.concept_id = ca1.descendant_concept_id
                    and c1.vocabulary_id = 'ATC'
                    and c1.concept_class_id = 'ATC 2nd'
```

```
                  inner join
                  @cdmSchema.dbo.concept c2
                  on ca1.ancestor_concept_id = c2.concept_id
                  and c2.vocabulary_id = 'ATC'
            and c2.concept_class_id = 'ATC 1st'
                  group by c1.concept_id, c1.concept_name
                  ) atc3_to_atc1
            on atc5_to_atc3.atc3_concept_id = atc3_to_atc1.atc3_concept_id

            left join @cdmSchema.dbo.concept atc1
             on atc3_to_atc1.atc1_concept_id = atc1.concept_id
      ) concept_hierarchy
      on hr1.concept_id = concept_hierarchy.concept_id
,
(select count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id = 1
) denom
;




/*
DRUG_ERA

--drilldown of when drug occurs relative to index
--graph:  scatterplot
--analysis_id: 1820
--x:  time (30-day increments)
--y:  %
*/


select hr1.cohort_definition_id,
      'First' as record_type,
      c1.concept_id,
      c1.concept_name,
      hr1.duration,
      hr1.count_value,
      case when t1.count_value > 0 then 1.0*hr1.count_value / t1.count_value else 0
end as pct_persons
from
(select cohort_definition_id,
      cast(stratum_1 as integer) as concept_id,
      cast(stratum_2 as integer)*30 as duration,
      count_value
from @resultsSchema.dbo.@heraclesResultsTable
```

```
where analysis_id in (1870)
and cohort_definition_id in (@cohortDefinitionId)
) hr1
inner join
(
	select cohort_definition_id,
		-1* cast(stratum_1 as integer)*30 as duration,
		sum(count_value) over (partition by cohort_definition_id order by -1*
cast(stratum_1 as integer)*30 asc) as count_value
	from
	@resultsSchema.dbo.@heraclesResultsTable
	where analysis_id in (1805)
	and cohort_definition_id in (@cohortDefinitionId)
	and cast(stratum_1 as integer) > 0

	union

	select hr1.cohort_definition_id,
		cast(hr1.stratum_1 as integer)*30 as duration,
		t1.count_value - sum(hr1.count_value) over (partition by
hr1.cohort_definition_id order by cast(hr1.stratum_1 as integer)*30 asc) as
count_value
	from
	@resultsSchema.dbo.@heraclesResultsTable hr1
	inner join
	(select cohort_definition_id, sum(count_value) as count_value
	from @resultsSchema.dbo.@heraclesResultsTable
	where analysis_id = 1806
	and cohort_definition_id in (@cohortDefinitionId)
	group by cohort_definition_id) t1
	on hr1.cohort_definition_id = t1.cohort_definition_id
	where hr1.analysis_id in (1806)
	and hr1.cohort_definition_id in (@cohortDefinitionId)

) t1
on hr1.cohort_definition_id = t1.cohort_definition_id
and hr1.duration = t1.duration
inner join
(select cohort_definition_id,
	cast(stratum_1 as integer) as concept_id,
	sum(count_value) as count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1870)
and cohort_definition_id in (@cohortDefinitionId)
group by cohort_definition_id,
	cast(stratum_1 as integer)
having sum(count_value) > @minCovariatePersonCount
) ct1
```

```
on hr1.cohort_definition_id = ct1.cohort_definition_id
and hr1.concept_id = ct1.concept_id
inner join
@cdmSchema.dbo.concept c1
on hr1.concept_id = c1.concept_id
where t1.count_value > @minIntervalPersonCount

union

select hr1.cohort_definition_id,
	'All' as record_type,
	c1.concept_id,
	c1.concept_name,
	hr1.duration,
	hr1.count_value,
	case when t1.count_value > 0 then 1.0*hr1.count_value / t1.count_value else 0
end as pct_persons
from
(select cohort_definition_id,
	cast(stratum_1 as integer) as concept_id,
	cast(stratum_2 as integer)*30 as duration,
	count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1871)
and cohort_definition_id in (@cohortDefinitionId)
) hr1
inner join
(
	select cohort_definition_id,
		-1* cast(stratum_1 as integer)*30 as duration,
		sum(count_value) over (partition by cohort_definition_id order by -1*
cast(stratum_1 as integer)*30 asc) as count_value
	from
	@resultsSchema.dbo.@heraclesResultsTable
	where analysis_id in (1805)
	and cohort_definition_id in (@cohortDefinitionId)
	and cast(stratum_1 as integer) > 0

	union

	select hr1.cohort_definition_id,
		cast(hr1.stratum_1 as integer)*30 as duration,
		t1.count_value - sum(hr1.count_value) over (partition by
hr1.cohort_definition_id order by cast(hr1.stratum_1 as integer)*30 asc) as
count_value
	from
	@resultsSchema.dbo.@heraclesResultsTable hr1
	inner join
```

```
	(select cohort_definition_id, sum(count_value) as count_value
	from @resultsSchema.dbo.@heraclesResultsTable
	where analysis_id = 1806
	and cohort_definition_id in (@cohortDefinitionId)
	group by cohort_definition_id) t1
	on hr1.cohort_definition_id = t1.cohort_definition_id
	where hr1.analysis_id in (1806)
	and hr1.cohort_definition_id in (@cohortDefinitionId)

) t1
on hr1.cohort_definition_id = t1.cohort_definition_id
and hr1.duration = t1.duration
inner join
(select cohort_definition_id,
	cast(stratum_1 as integer) as concept_id,
	sum(count_value) as count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1871)
and cohort_definition_id in (@cohortDefinitionId)
group by cohort_definition_id,
	cast(stratum_1 as integer)
having sum(count_value) > @minCovariatePersonCount
) ct1
on hr1.cohort_definition_id = ct1.cohort_definition_id
and hr1.concept_id = ct1.concept_id
inner join
@cdmSchema.dbo.concept c1
on hr1.concept_id = c1.concept_id
where t1.count_value > @minIntervalPersonCount
;







/*
PROCEDURE_OCCURRENCE

--treemap of all conditions
--analysis_id: 1870
--size - prevalence of drug
--color:  risk difference of prevalence before / after index
*/
```

```
select   concept_hierarchy.concept_id,
       isNull(concept_hierarchy.proc_concept_name,'NA') as concept_name,
       isNull(concept_hierarchy.level2_concept_name,'NA') as level2_concept_name,
       isNull(concept_hierarchy.level3_concept_name,'NA') as level3_concept_name,
       isNull(concept_hierarchy.level4_concept_name,'NA') as level4_concept_name,
        isNull(concept_hierarchy.level4_concept_name,'NA')
       + '||' + isNull(concept_hierarchy.level3_concept_name,'NA')
       + '||' + isNull(concept_hierarchy.level2_concept_name,'NA')
       + '||' + isNull(concept_hierarchy.proc_concept_name,'NA') as concept_path,
       1.0*hr1.num_persons / denom.count_value as percent_persons,
       1.0*hr1.num_persons_before / denom.count_value as percent_persons_before,
       1.0*hr1.num_persons_after / denom.count_value as percent_persons_after,
       1.0*(hr1.num_persons_after - hr1.num_persons_before)/denom.count_value as
risk_diff_after_before,
       log(1.0*(hr1.num_persons_after + 0.5) / (hr1.num_persons_before + 0.5)) as
logRR_after_before
from
(select cast(stratum_1 as integer) as concept_id,
       sum(count_value) as num_persons,
       sum(case when stratum_2 < 0 then count_value else 0 end) as
num_persons_before,
       sum(case when stratum_2 > 0 then count_value else 0 end) as num_persons_after
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1830) --first occurrence of procedure
and cohort_definition_id in (@cohortDefinitionId)
group by cast(stratum_1 as int)
) hr1
inner join
       (
              select procs.concept_id,
              procs.proc_concept_name,
              max(proc_hierarchy.os3_concept_name) as level2_concept_name,
              max(proc_hierarchy.os2_concept_name) as level3_concept_name,
              max(proc_hierarchy.os1_concept_name) as level4_concept_name
        from
              (
              select c1.concept_id,
                     v1.vocabulary_name + ' ' + c1.concept_code + ': ' +
c1.concept_name as proc_concept_name
              from @cdmSchema.dbo.concept c1
                     inner join @cdmSchema.dbo.vocabulary v1
                     on c1.vocabulary_id = v1.vocabulary_id
              where (
                     c1.vocabulary_id in ('ICD9Proc', 'HCPCS','CPT4')
                     or (c1.vocabulary_id = 'SNOMED' and c1.concept_class_id =
'Procedure')
                     )
              ) procs
```

```
      left join
            (select ca0.DESCENDANT_CONCEPT_ID, max(ca0.ancestor_concept_id) as
ancestor_concept_id
            from @cdmSchema.dbo.concept_ancestor ca0
            inner join
            (select distinct c2.concept_id as os3_concept_id
             from @cdmSchema.dbo.concept_ancestor ca1
                  inner join
                  @cdmSchema.dbo.concept c1
                  on ca1.DESCENDANT_CONCEPT_ID = c1.concept_id
                  inner join
                  @cdmSchema.dbo.concept_ancestor ca2
                  on c1.concept_id = ca2.ANCESTOR_CONCEPT_ID
                  inner join
                  @cdmSchema.dbo.concept c2
                  on ca2.DESCENDANT_CONCEPT_ID = c2.concept_id
             where ca1.ancestor_concept_id = 4040390
             and ca1.Min_LEVELS_OF_SEPARATION = 2
             and ca2.MIN_LEVELS_OF_SEPARATION = 1
        ) t1

            on ca0.ANCESTOR_CONCEPT_ID = t1.os3_concept_id

            group by ca0.descendant_concept_id

            ) ca1
            on procs.concept_id = ca1.DESCENDANT_CONCEPT_ID
      left join
      (
       select proc_by_os1.os1_concept_name,
            proc_by_os2.os2_concept_name,
            proc_by_os3.os3_concept_name,
            proc_by_os3.os3_concept_id
      from
       (select DESCENDANT_CONCEPT_ID as os1_concept_id, concept_name as
os1_concept_name
       from @cdmSchema.dbo.concept_ancestor ca1
            inner join
            @cdmSchema.dbo.concept c1
            on ca1.DESCENDANT_CONCEPT_ID = c1.concept_id
       where ancestor_concept_id = 4040390
       and Min_LEVELS_OF_SEPARATION = 1
       ) proc_by_os1

       inner join
       (select max(c1.CONCEPT_ID) as os1_concept_id, c2.concept_id as
os2_concept_id, c2.concept_name as os2_concept_name
```

```
        from @cdmSchema.dbo.concept_ancestor ca1
              inner join
              @cdmSchema.dbo.concept c1
              on ca1.DESCENDANT_CONCEPT_ID = c1.concept_id
              inner join
              @cdmSchema.dbo.concept_ancestor ca2
              on c1.concept_id = ca2.ANCESTOR_CONCEPT_ID
              inner join
              @cdmSchema.dbo.concept c2
              on ca2.DESCENDANT_CONCEPT_ID = c2.concept_id
        where ca1.ancestor_concept_id = 4040390
        and ca1.Min_LEVELS_OF_SEPARATION = 1
        and ca2.MIN_LEVELS_OF_SEPARATION = 1
        group by c2.concept_id, c2.concept_name
        ) proc_by_os2
        on proc_by_os1.os1_concept_id = proc_by_os2.os1_concept_id

        inner join
        (select max(c1.CONCEPT_ID) as os2_concept_id, c2.concept_id as
os3_concept_id, c2.concept_name as os3_concept_name
        from @cdmSchema.dbo.concept_ancestor ca1
              inner join
              @cdmSchema.dbo.concept c1
              on ca1.DESCENDANT_CONCEPT_ID = c1.concept_id
              inner join
              @cdmSchema.dbo.concept_ancestor ca2
              on c1.concept_id = ca2.ANCESTOR_CONCEPT_ID
              inner join
              @cdmSchema.dbo.concept c2
              on ca2.DESCENDANT_CONCEPT_ID = c2.concept_id
        where ca1.ancestor_concept_id = 4040390
        and ca1.Min_LEVELS_OF_SEPARATION = 2
        and ca2.MIN_LEVELS_OF_SEPARATION = 1
         group by c2.concept_id, c2.concept_name
        ) proc_by_os3
        on proc_by_os2.os2_concept_id = proc_by_os3.os2_concept_id
       ) proc_hierarchy
       on ca1.ancestor_concept_id = proc_hierarchy.os3_concept_id
       group by procs.concept_id,
              procs.proc_concept_name

       ) concept_hierarchy
       on hr1.concept_id = concept_hierarchy.concept_id
,
(select count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id = 1
) denom
```

```
;



/*
PROCEDURE_OCURRENCE

--drilldown of when procedure occurs relative to index
--graph:  scatterplot
--analysis_id: 1820
--x:  time (30-day increments)
--y:  %
*/


select hr1.cohort_definition_id,
       'First' as record_type,
       c1.concept_id,
       c1.concept_name,
       hr1.duration,
       hr1.count_value,
       case when t1.count_value > 0 then 1.0*hr1.count_value / t1.count_value else 0
end as pct_persons
from
(select cohort_definition_id,
       cast(stratum_1 as integer) as concept_id,
       cast(stratum_2 as integer)*30 as duration,
       count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1830)
and cohort_definition_id in (@cohortDefinitionId)
) hr1
inner join
(
       select cohort_definition_id,
              -1* cast(stratum_1 as integer)*30 as duration,
              sum(count_value) over (partition by cohort_definition_id order by -1*
cast(stratum_1 as integer)*30 asc) as count_value
       from
       @resultsSchema.dbo.@heraclesResultsTable
       where analysis_id in (1805)
       and cohort_definition_id in (@cohortDefinitionId)
       and cast(stratum_1 as integer) > 0

       union

       select hr1.cohort_definition_id,
```

```
            cast(hr1.stratum_1 as integer)*30 as duration,
            t1.count_value - sum(hr1.count_value) over (partition by
hr1.cohort_definition_id order by cast(hr1.stratum_1 as integer)*30 asc) as
count_value
      from
      @resultsSchema.dbo.@heraclesResultsTable hr1
      inner join
      (select cohort_definition_id, sum(count_value) as count_value
      from @resultsSchema.dbo.@heraclesResultsTable
      where analysis_id = 1806
      and cohort_definition_id in (@cohortDefinitionId)
      group by cohort_definition_id) t1
      on hr1.cohort_definition_id = t1.cohort_definition_id
      where hr1.analysis_id in (1806)
      and hr1.cohort_definition_id in (@cohortDefinitionId)

) t1
on hr1.cohort_definition_id = t1.cohort_definition_id
and hr1.duration = t1.duration
inner join
(select cohort_definition_id,
      cast(stratum_1 as integer) as concept_id,
      sum(count_value) as count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1830)
and cohort_definition_id in (@cohortDefinitionId)
group by cohort_definition_id,
      cast(stratum_1 as integer)
having sum(count_value) > @minCovariatePersonCount
) ct1
on hr1.cohort_definition_id = ct1.cohort_definition_id
and hr1.concept_id = ct1.concept_id
inner join
@cdmSchema.dbo.concept c1
on hr1.concept_id = c1.concept_id
where t1.count_value > @minIntervalPersonCount

union

select hr1.cohort_definition_id,
      'All' as record_type,
      c1.concept_id,
      c1.concept_name,
      hr1.duration,
      hr1.count_value,
      case when t1.count_value > 0 then 1.0*hr1.count_value / t1.count_value else 0
end as pct_persons
from
```

```
(select cohort_definition_id,
       cast(stratum_1 as integer) as concept_id,
       cast(stratum_2 as integer)*30 as duration,
       count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1831)
and cohort_definition_id in (@cohortDefinitionId)
) hr1
inner join
(
       select cohort_definition_id,
              -1* cast(stratum_1 as integer)*30 as duration,
              sum(count_value) over (partition by cohort_definition_id order by -1*
cast(stratum_1 as integer)*30 asc) as count_value
       from
       @resultsSchema.dbo.@heraclesResultsTable
       where analysis_id in (1805)
       and cohort_definition_id in (@cohortDefinitionId)
       and cast(stratum_1 as integer) > 0

       union

       select hr1.cohort_definition_id,
              cast(hr1.stratum_1 as integer)*30 as duration,
              t1.count_value - sum(hr1.count_value) over (partition by
hr1.cohort_definition_id order by cast(hr1.stratum_1 as integer)*30 asc) as
count_value
       from
       @resultsSchema.dbo.@heraclesResultsTable hr1
       inner join
       (select cohort_definition_id, sum(count_value) as count_value
       from @resultsSchema.dbo.@heraclesResultsTable
       where analysis_id = 1806
       and cohort_definition_id in (@cohortDefinitionId)
       group by cohort_definition_id) t1
       on hr1.cohort_definition_id = t1.cohort_definition_id
       where hr1.analysis_id in (1806)
       and hr1.cohort_definition_id in (@cohortDefinitionId)

) t1
on hr1.cohort_definition_id = t1.cohort_definition_id
and hr1.duration = t1.duration
inner join
(select cohort_definition_id,
       cast(stratum_1 as integer) as concept_id,
       sum(count_value) as count_value
from @resultsSchema.dbo.@heraclesResultsTable
where analysis_id in (1831)
```

```
and cohort_definition_id in (@cohortDefinitionId)
group by cohort_definition_id,
       cast(stratum_1 as integer)
having sum(count_value) > @minCovariatePersonCount
) ct1
on hr1.cohort_definition_id = ct1.cohort_definition_id
and hr1.concept_id = ct1.concept_id
inner join
@cdmSchema.dbo.concept c1
on hr1.concept_id = c1.concept_id
where t1.count_value > @minIntervalPersonCount
;
```