



Estimation of prescription duration from dispensing dates – challenges and new opportunities with the Waiting Time Distribution

Henrik Støvring

stovring@biomed.au.dk

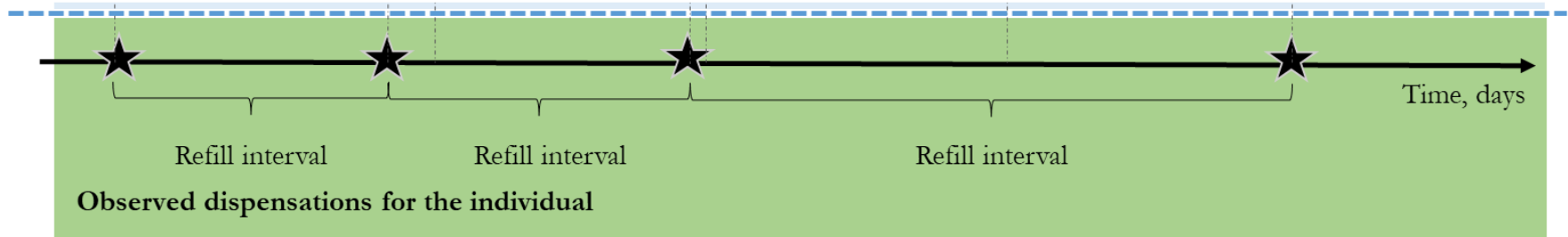
Based on joint work with
Sabrina Giometto, Malcolm Gillies, Katrine Bødkergaard, Anton
Pottegård, Jesper Hallas and others
May 4, 2026 – DSTS, Aalborg

Overview

1. Basic challenge: treatment episode
2. Introduction to the WTD (reverse WTD)
3. Random index dates and statistical efficiency
4. Saturated sampling and statistical efficiency

1. Basic challenge: treatment episodes

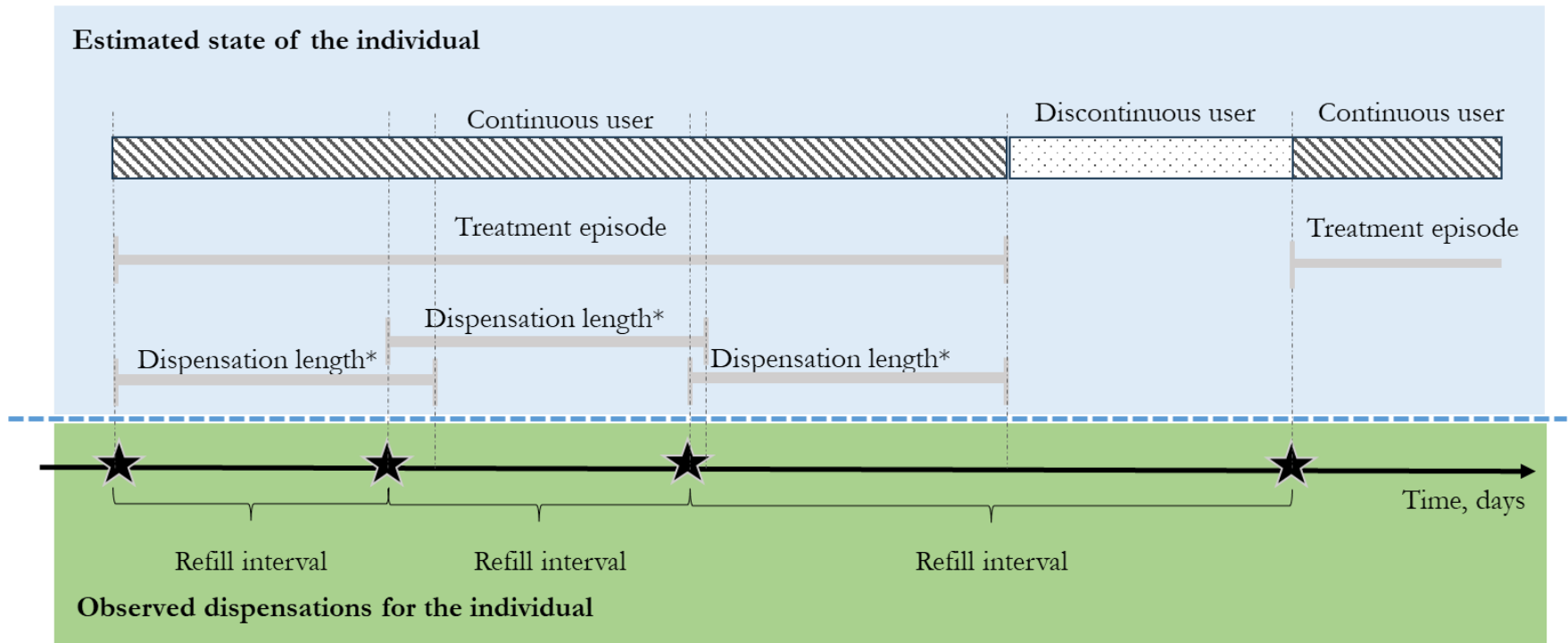
Pharmacy dispensing to treatment episodes



★ Dispensation of medication at pharmacy

* Dispensation length estimated based on percentiles of the interarrival distribution.

Pharmacy dispensing to treatment episodes



★ Dispensation of medication at pharmacy

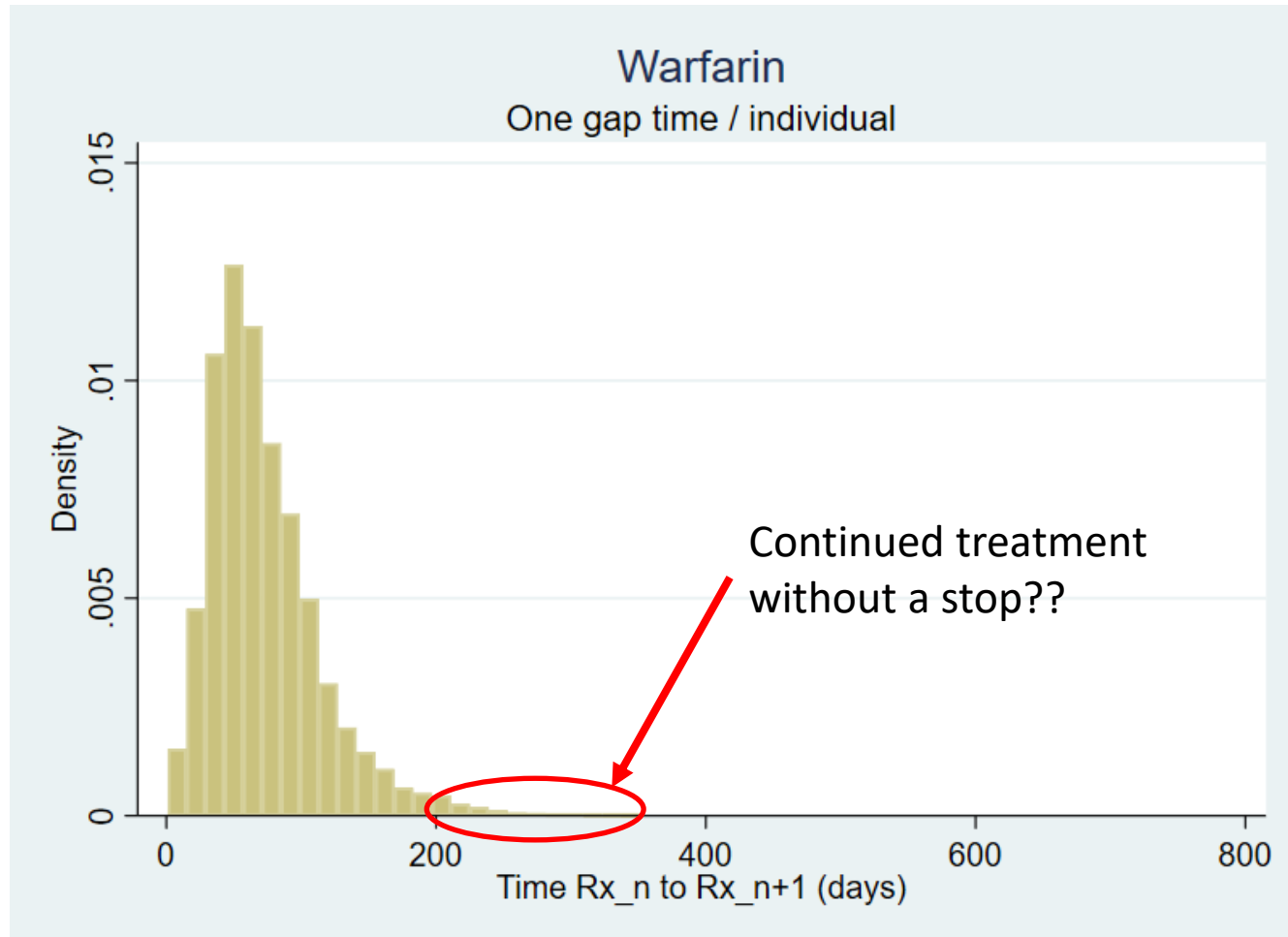
* Dispensation length estimated based on percentiles of the interarrival distribution.

Standard decision rules

- Number of DDDs (defined daily doses) in obtained medication:
 - Patient takes one DDD per day
→ number of DDDs = number of days treated after prescription redemption
 - Patient takes half a DDD per day
→ $2 \times$ number of DDDs = number of days treated after prescription redemption
- Number of pills in obtained medication
 - (same principle as above)
- Note: Often a grace period is added, and non-used medication may be pasted to subsequent prescriptions
- **Main limitations of decision rules**
 - Assume that all patients are treated with equal dose
 - Ignores observed patterns in prescription redemptions

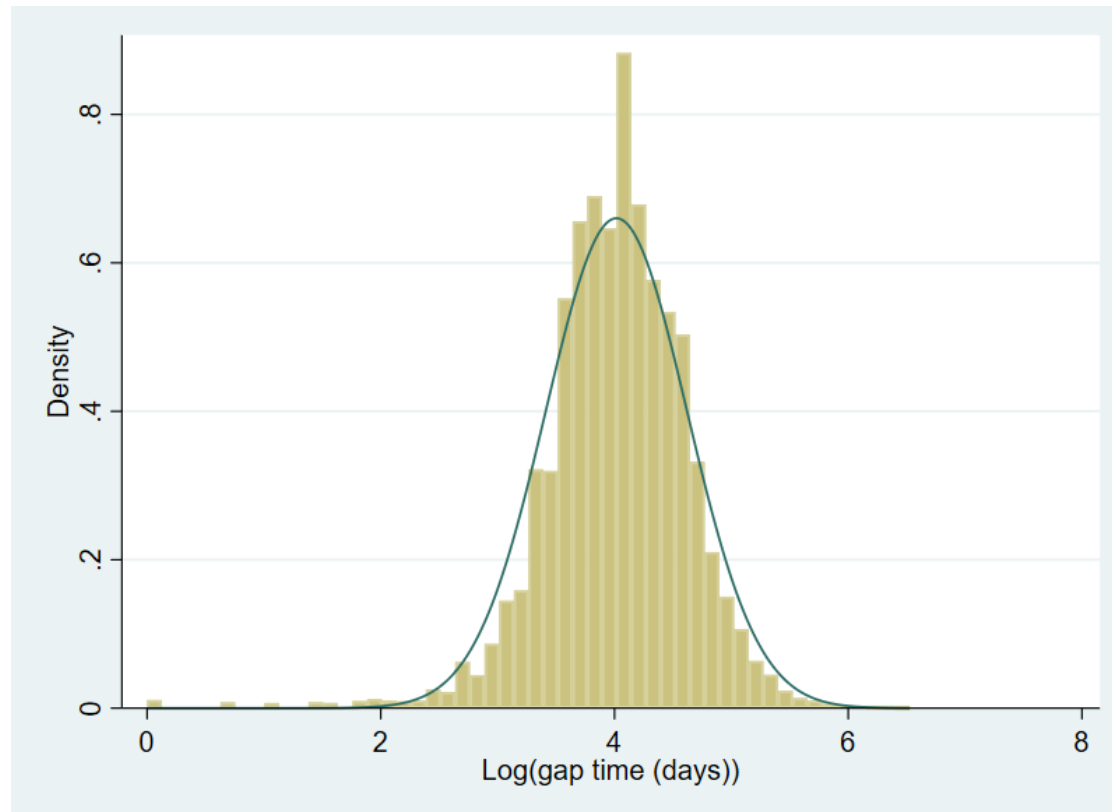
One random gap time per patient, OPED 2013

- Median 64 days, $n = 19,446$ gap times
- Range 1 to 349 days



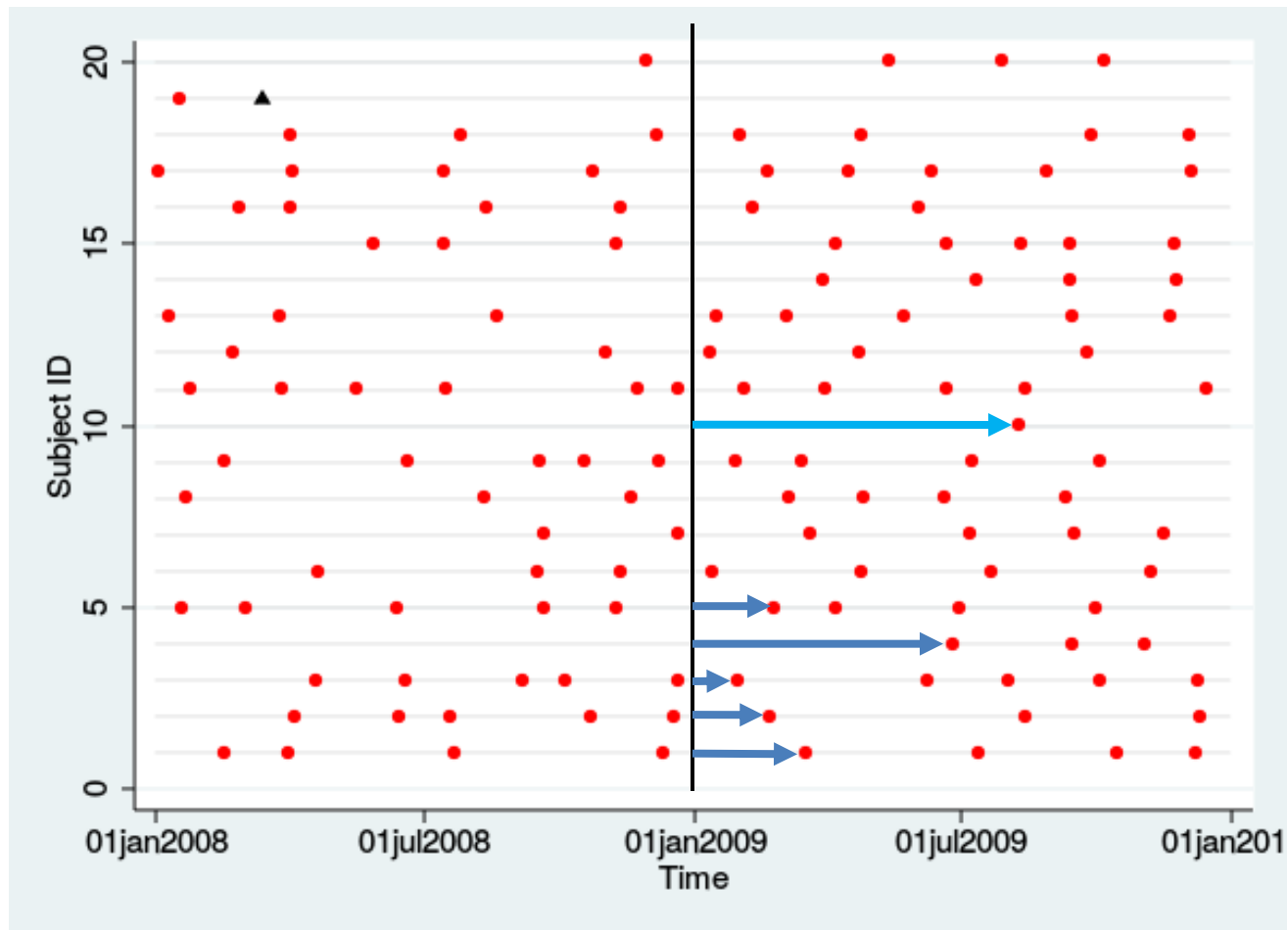
A first attempt at a parametric analysis

- One randomly selected gap time for all patients, OPED 2013-2014
- Median of 62.6 days estimated from log-normal

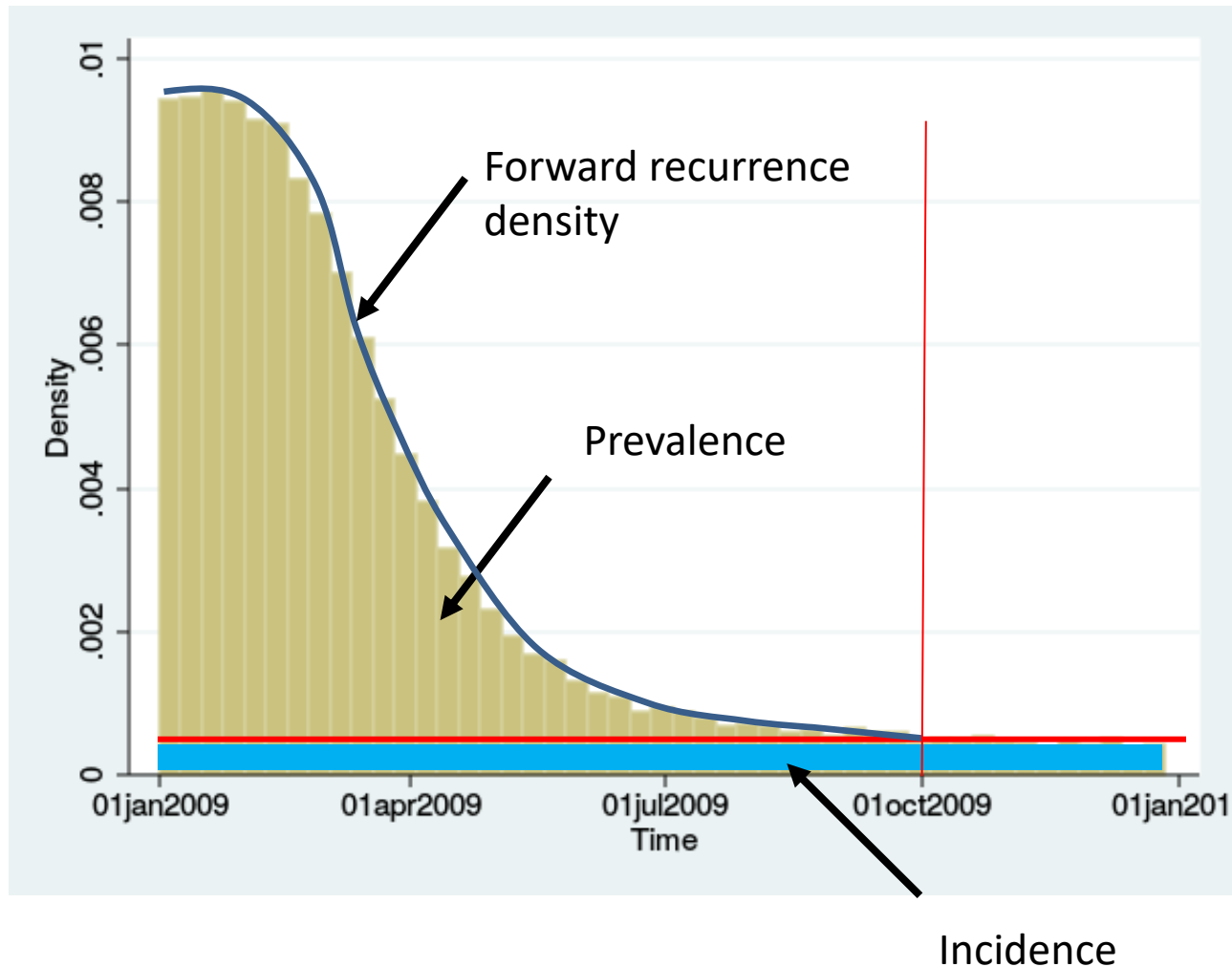


2. (Reverse) Waiting Time Distribution (rWTD)

Construction of Waiting Time Distribution



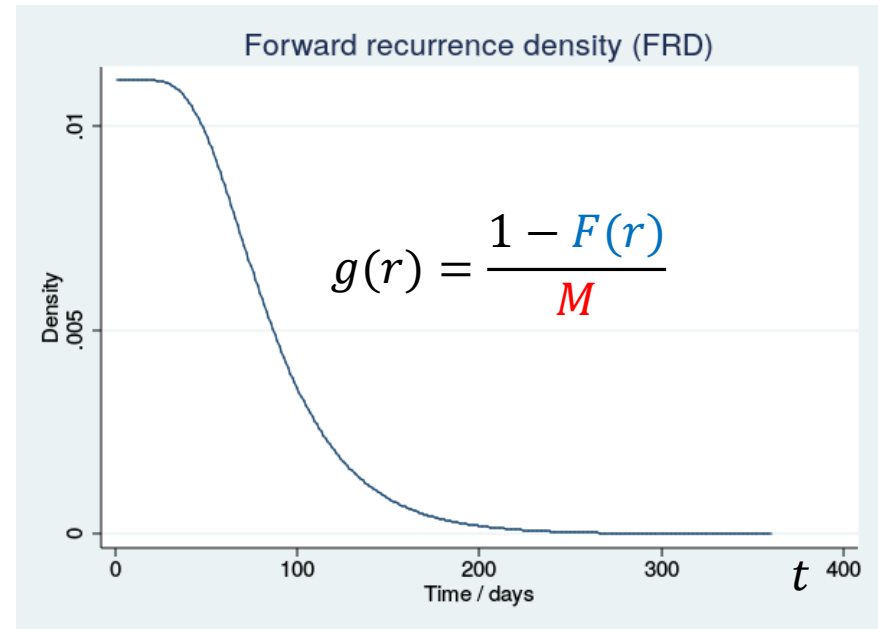
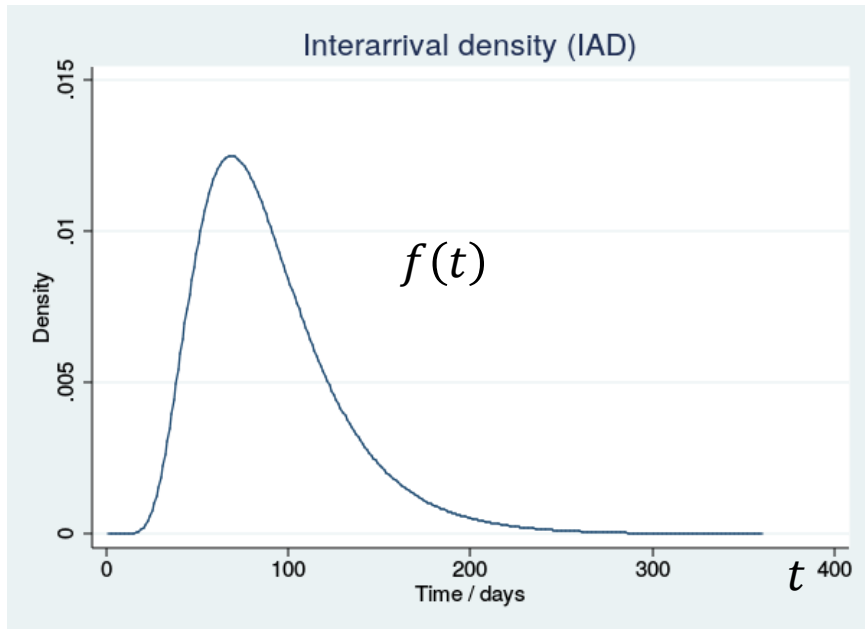
Components of the waiting time distribution



First suggested by *Hallas et al* (1997)

How length bias works

- Links distribution of time from R_{x_k} to $R_{x_{k+1}}$ (IAD) to distribution from index date to first subsequent R_k (FRD)



Where:

$F(t)$: the cumulative distribution function of the IAD

M : the mean of the IAD

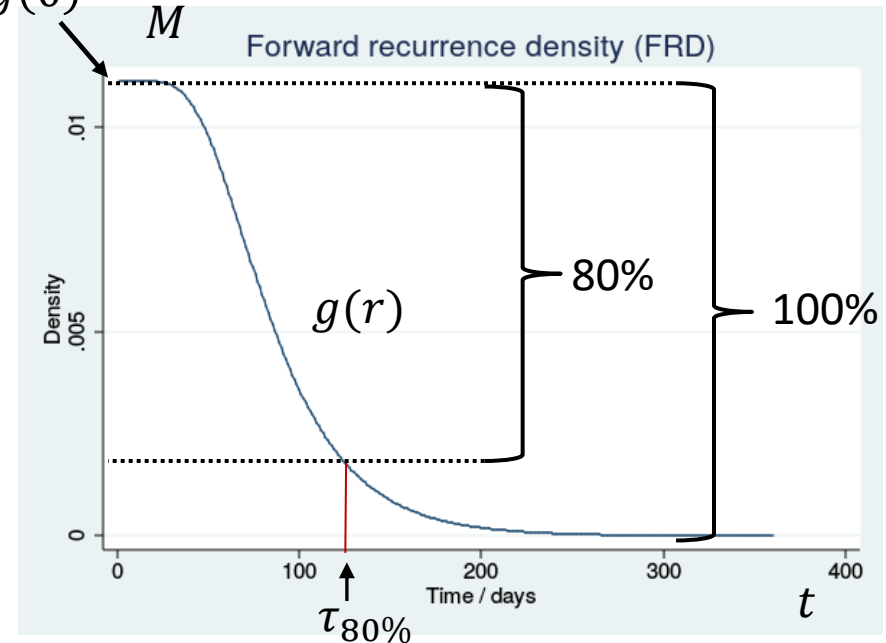
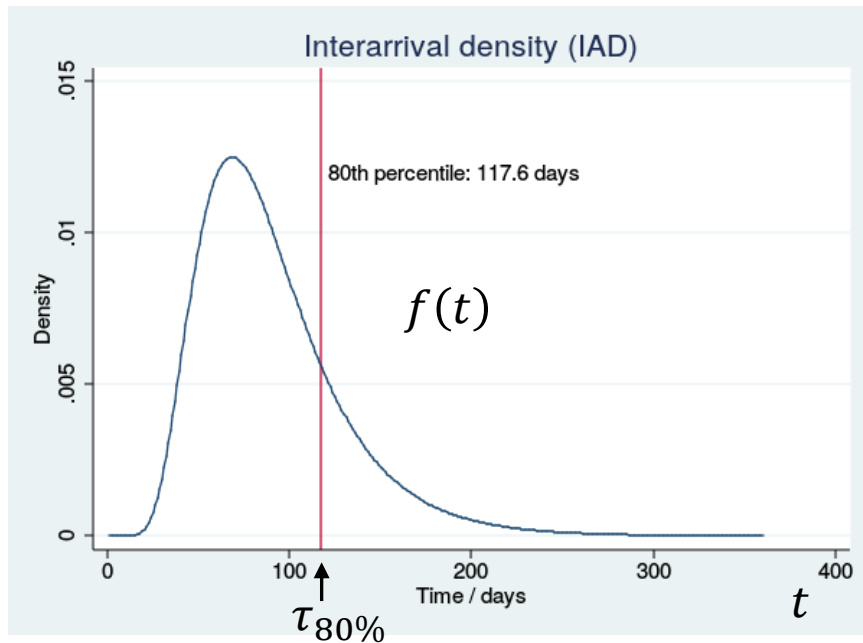
Percentiles in IAD and FRD

- Suppose we can estimate this WTD model (r is time to first Rx, γ is fraction of prevalents among patients in WTD, contribution for one individual):

$$L(r; \gamma, \theta) = \gamma g(r; \theta) + (1 - \gamma) \frac{1}{\delta}$$

$$g(r) = \frac{1 - F(r)}{M}$$

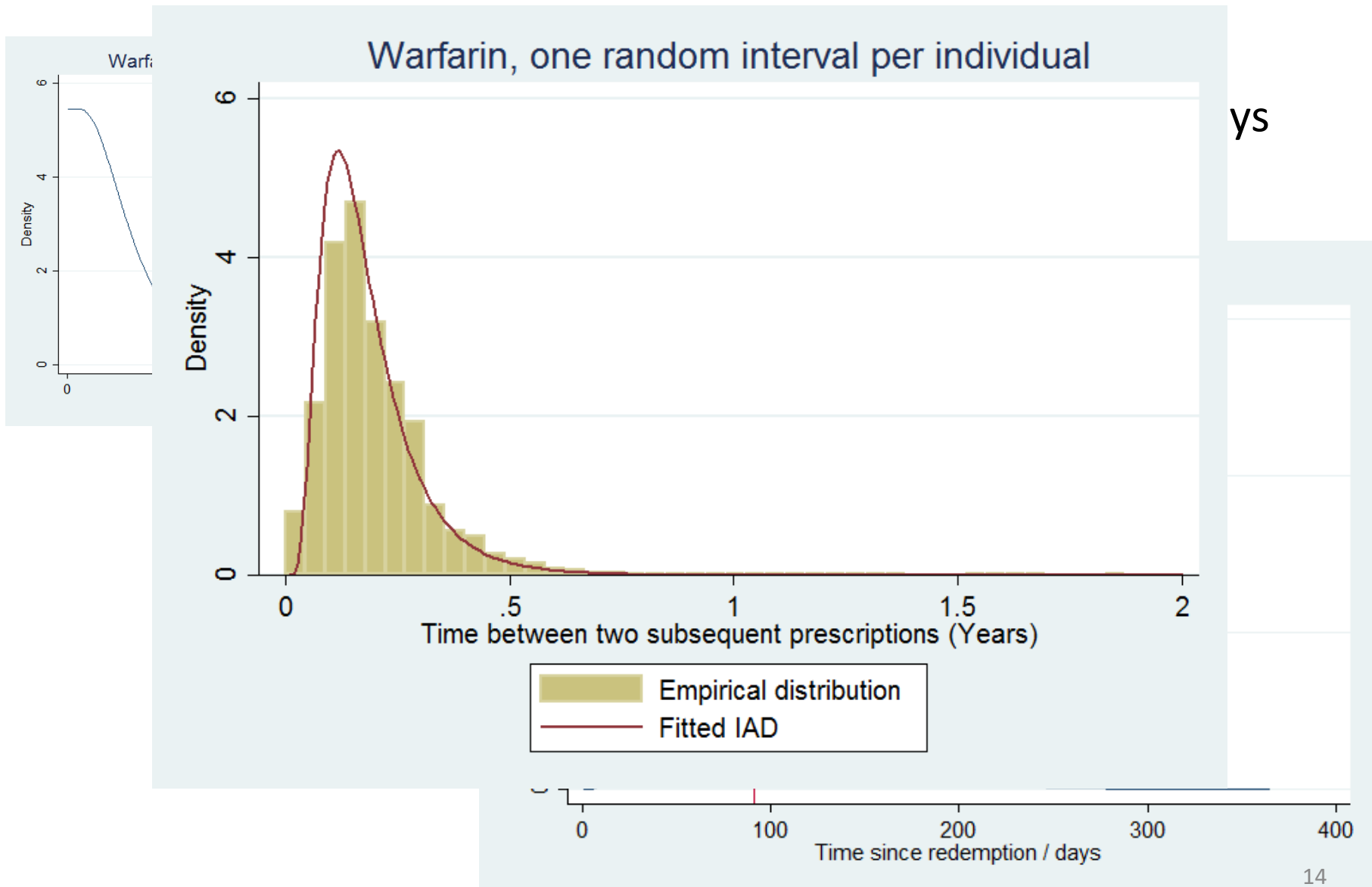
$$g(0) = \frac{1}{M}$$



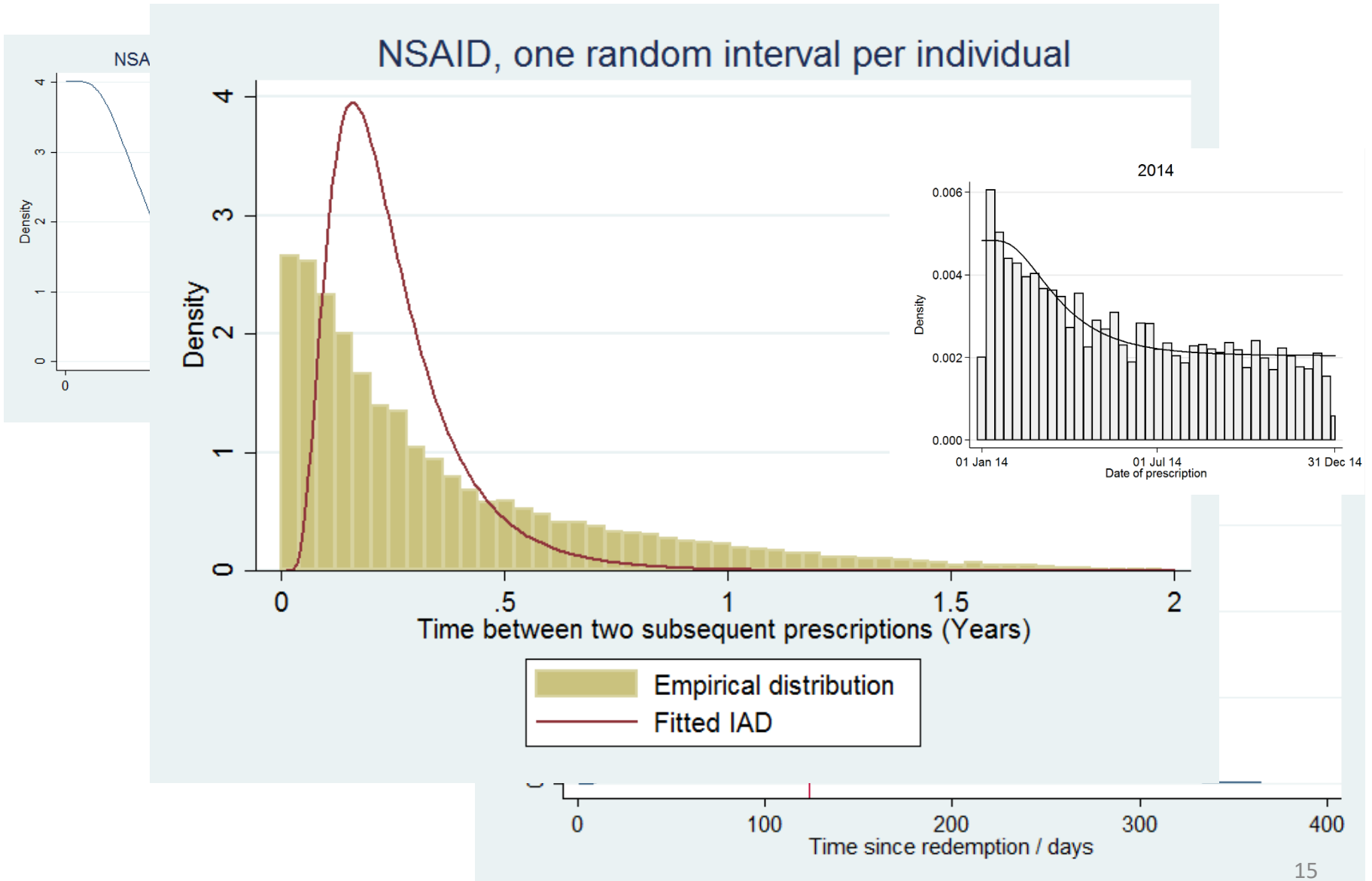
$\tau_{80\%}$ can be found from $g(r)$ as the r , which satisfies

$$M \cdot g(r) = (100\% - 80\%)$$

Does it work? – The poster boy: Warfarin

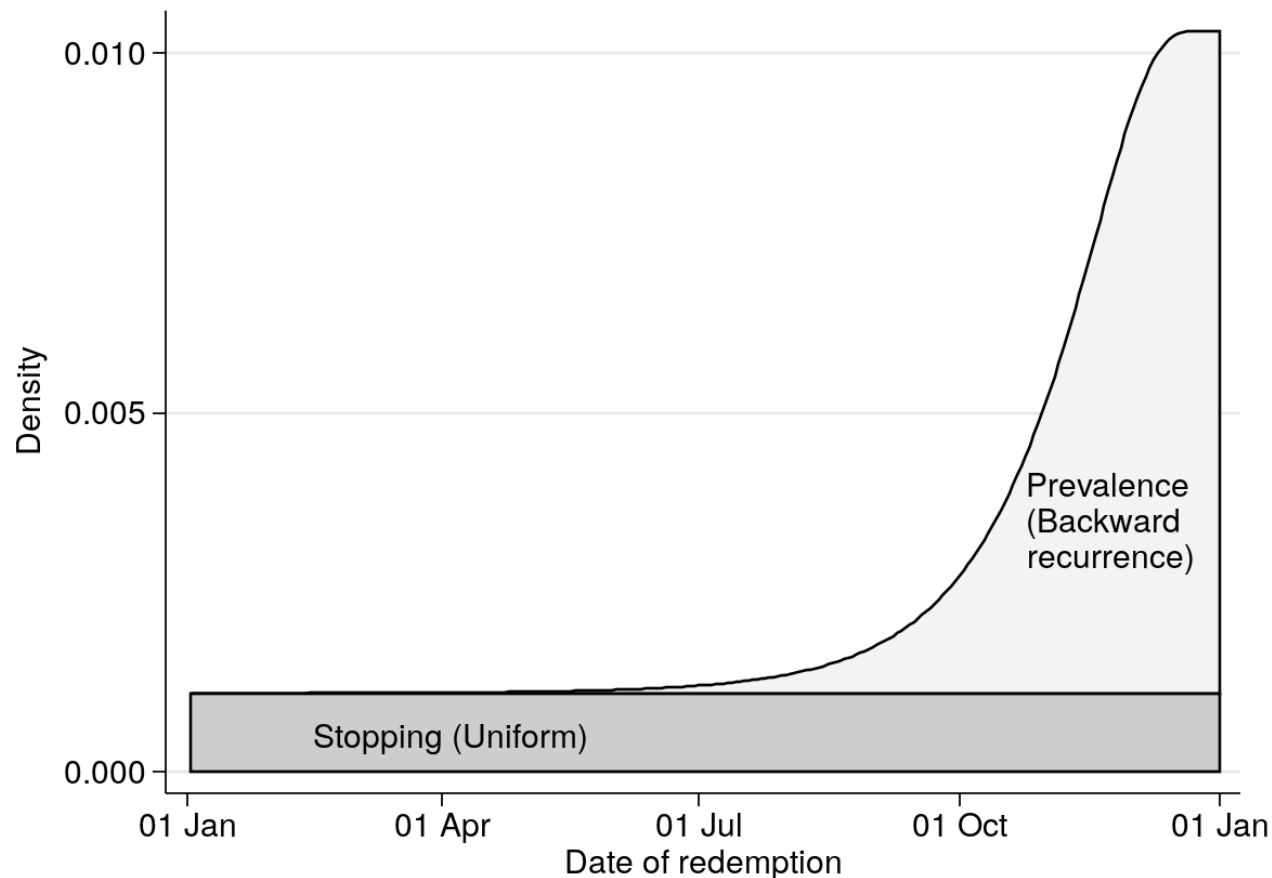


Does it work? – The black sheep: NSAIDs



Reverse WTD

- **Answers the questions:** What is the distribution of time from the last prescription in an interval to the end of the interval?
- Backward recurrence distribution also allows estimation of IAD
- **Allows regression-like analysis**



Warfarin: Example of estimation with WTD

```
. keep if year(ekspeditionsdato) == 2013
```

```
. bysort projectkey (ekspeditionsdato): keep if _n == _N
```

```
. wtdttt ekspeditionsdato, reverse dist(lnorm) start(1jan2013) end(31dec2013)
```

```
initial:      log likelihood = -135948.91
```

```
alternative:    log likelihood = -120905.19
```

<snip>

Iteration 7: log likelihood = -113184.8

Log likelihood = -113184.8

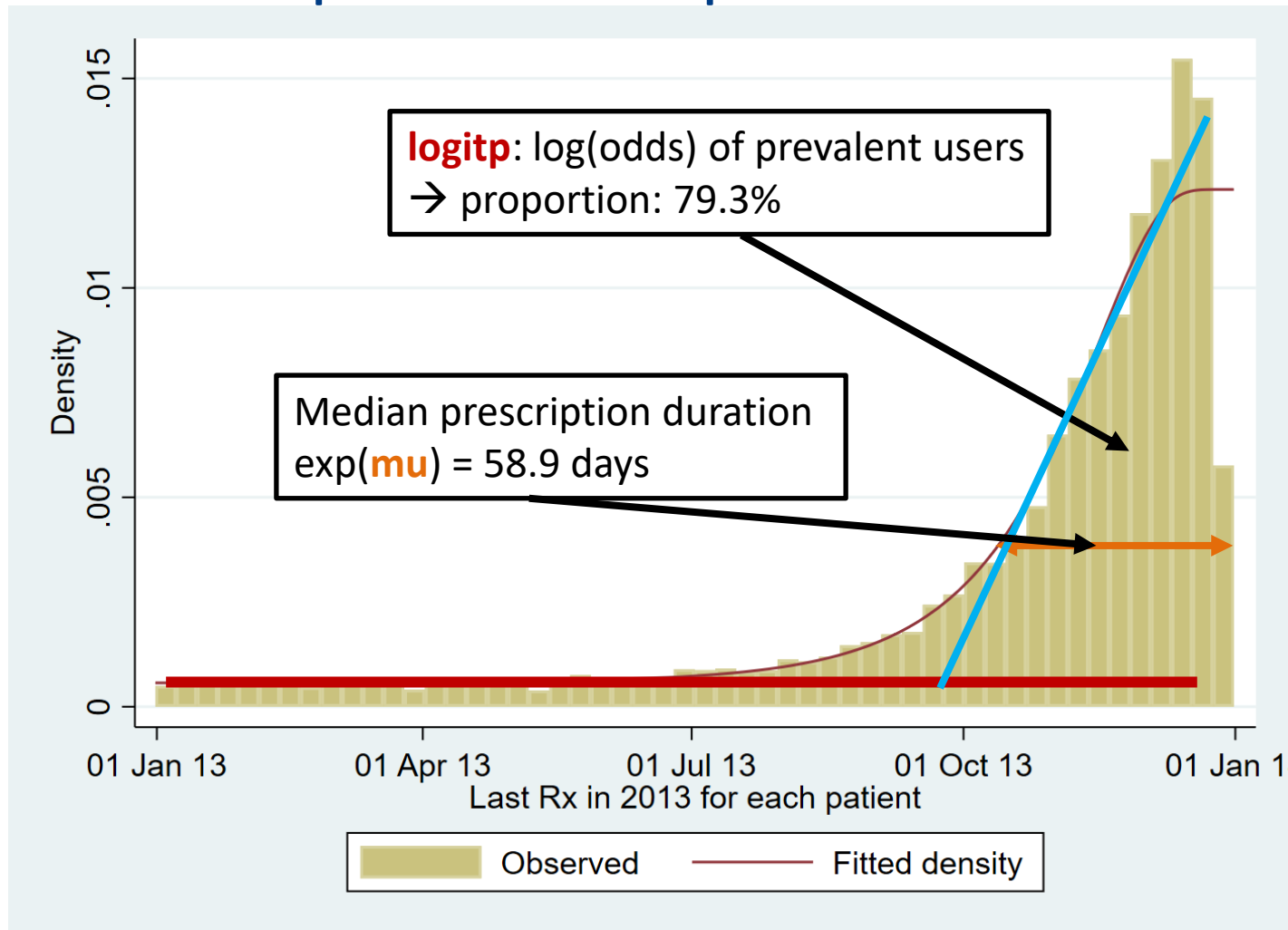
Number of obs = 21,675

Wald chi2(0) = .

```
Prob > chi2      =      .
```

ekspeditio~o	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
logitp						
_cons	1.34291	.0289142	46.44	0.000	1.286239	1.399581
mu						
_cons	4.0759	.011773	346.21	0.000	4.052826	4.098975
lnsigma						
_cons	-.661076	.0178061	-37.13	0.000	-.6959753	-.6261768

Warfarin: Interpretation of parameters

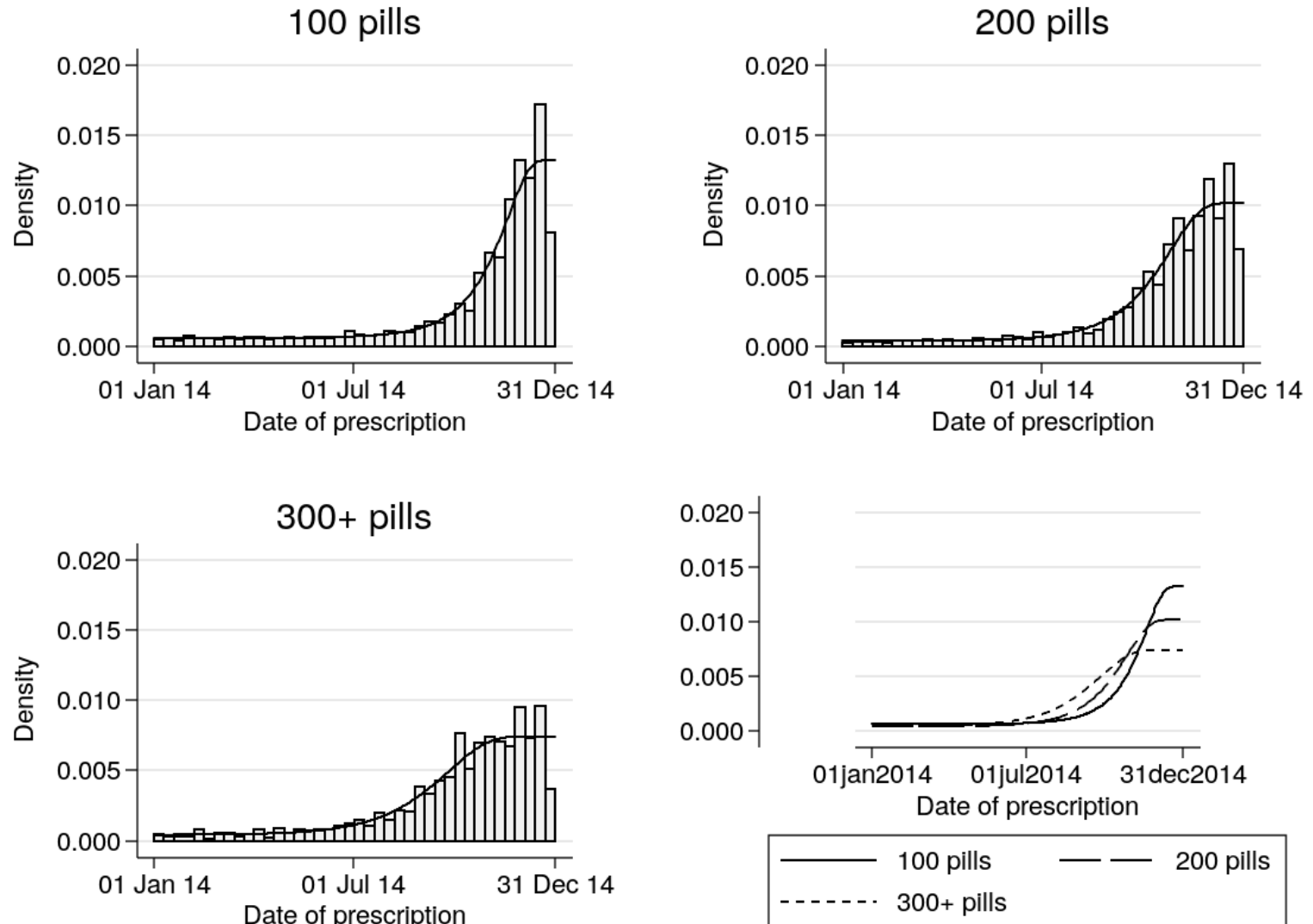


mu: log(Median interarrival time) – controls “width” of FRD

Insigma: SD(log interarrival time) – controls “steepness” of FRD

Reverse WTD with covariates

- **Warfarin**: Package size clearly impacts the reverse WTD



Reverse WTD with covariates – Warfarin

- Joint analysis – with number of pills as covariate
- Adjusted for
 - Sex
 - Age continuous (centered at age 50 years)
 - Interaction of sex and age

Number of pills	Odds ratio / odds (prevalence fraction)	Median ratio / Median (days)	σ ratio / σ (log-time)
100	-	-	-
200	1.69 (1.39; 2.05)	1.54 (1.46; 1.63)	0.85 (0.76; 0.94)
300+	1.54 (1.01; 2.25)	2.21 (2.02; 2.42)	0.77 (0.65; 0.92)
Reference (male, 50 years, 100 pills)	3.02 (2.55; 3.62)	39.7 (36.6; 43.1)	0.54 (0.46; 0.64)

Reverse WTD with covariates – R-package: wtdr

The screenshot shows the GitHub repository page for `mbg-unsw/wtdr`. The browser address bar shows `github.com/mbg-unsw/wtdr`. The repository page includes a file list, a README section, and a list of contributors.

File	Description	Time
README.md	Renamed package from 'Rwtdttt' t...	11 months ago
wtdr.Rproj	Renamed Rproj and pdf	11 months ago
wtdr.pdf	Closes #39	3 days ago

Contributors 3

-  **sabrinagiometto** Sabrina Giometto
-  **mbg-unsw** Malcolm Gillies
-  **HenrikStovring** Henrik Støving

Languages

R 100.0%

README  GPL-3.0 license

wtdr

wtdr is an R package to estimate parameters of the ordinary or reverse waiting time distribution (WTD) for applications in pharmacoepidemiology and drug utilisation research.

The package was previous named Rwtdttt and renamed in May 2025.

Installation

You can install `wtdr` from `github` using the `devtools` package

```
library(devtools)
install_github('mbg-unsw/wtdr')
```

3. Random index dates

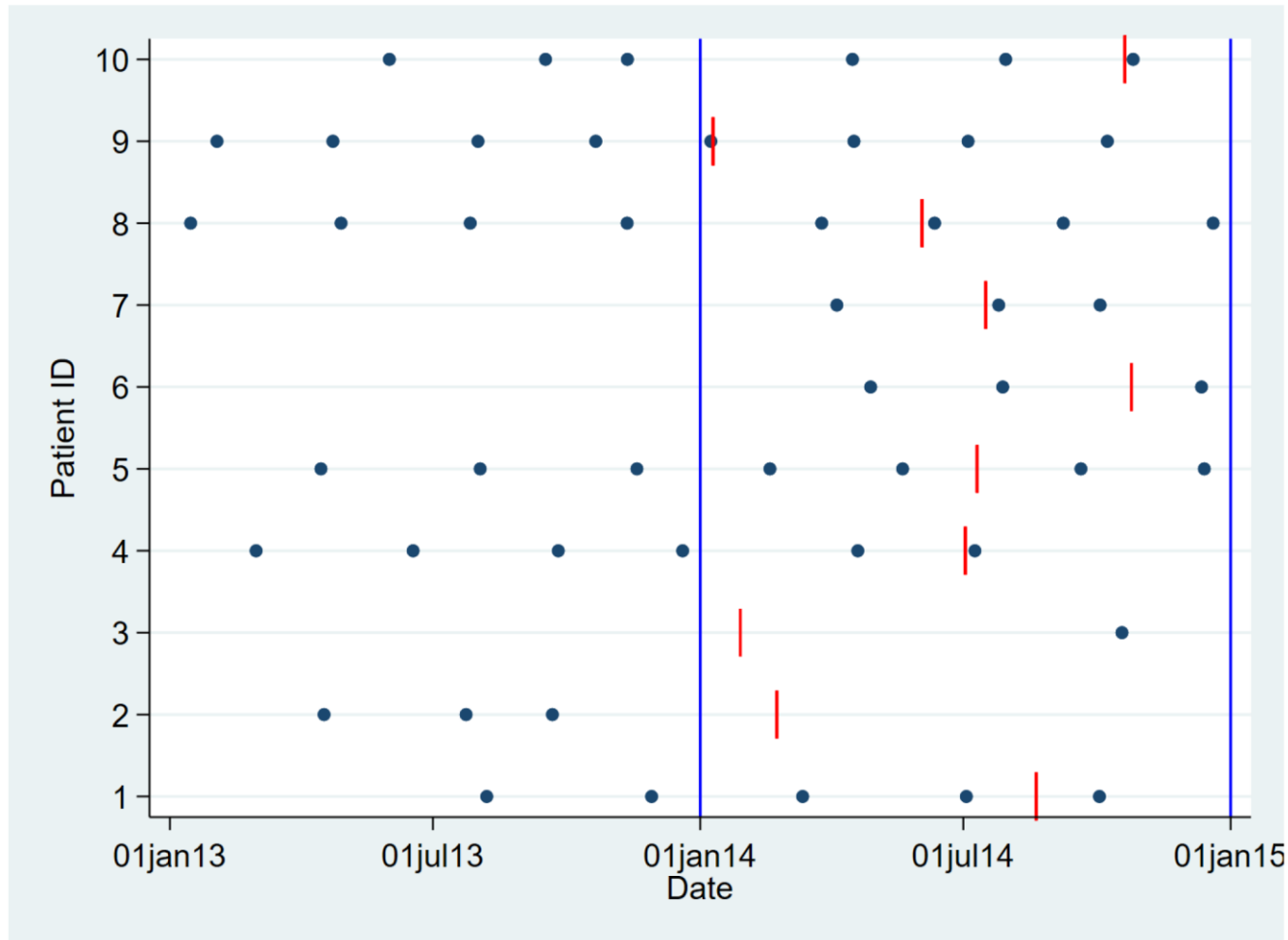
Problems of fixed index dates

- Requirement from renewal process theory:
Index date should be independent of process
- Violated if there is
 - seasonal variation
 - stockpiling
- Example: Norwegian prescriptions of warfarin (ATC B01AA03), 2013-2015

n	WTD	Index date	Estimated 80% percentile (days)
69,207	ordinary	Jan 1 2015	161.1 (159.7; 162.5)
77,674	reverse	Jan 1 2015	100.8 (99.7; 101.9)
79,850	reverse	Nov 1 2014	121.7 (120.6; 122.9)

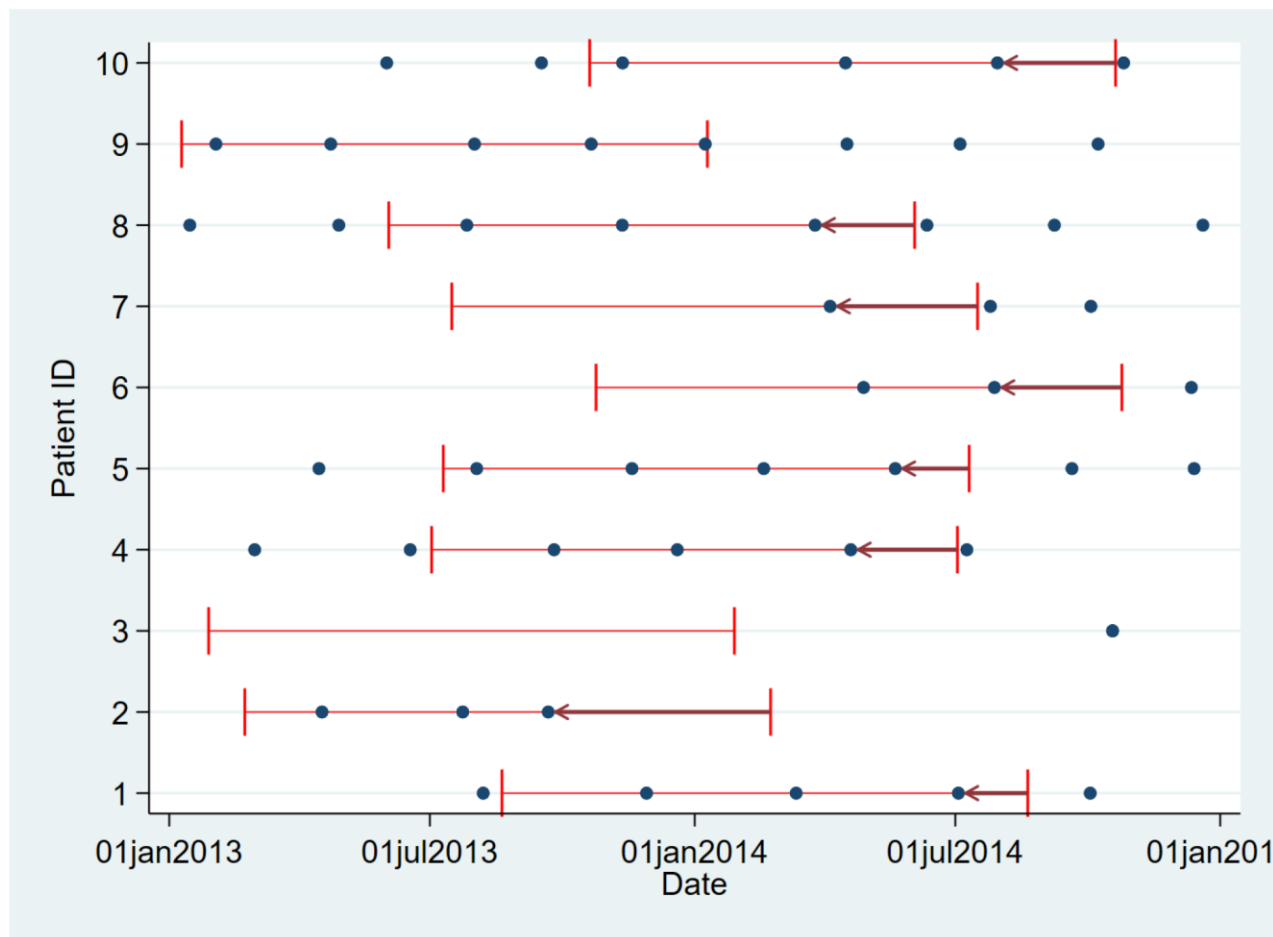
Idea: individual random index dates

- Step I: Sample individual index dates in interval



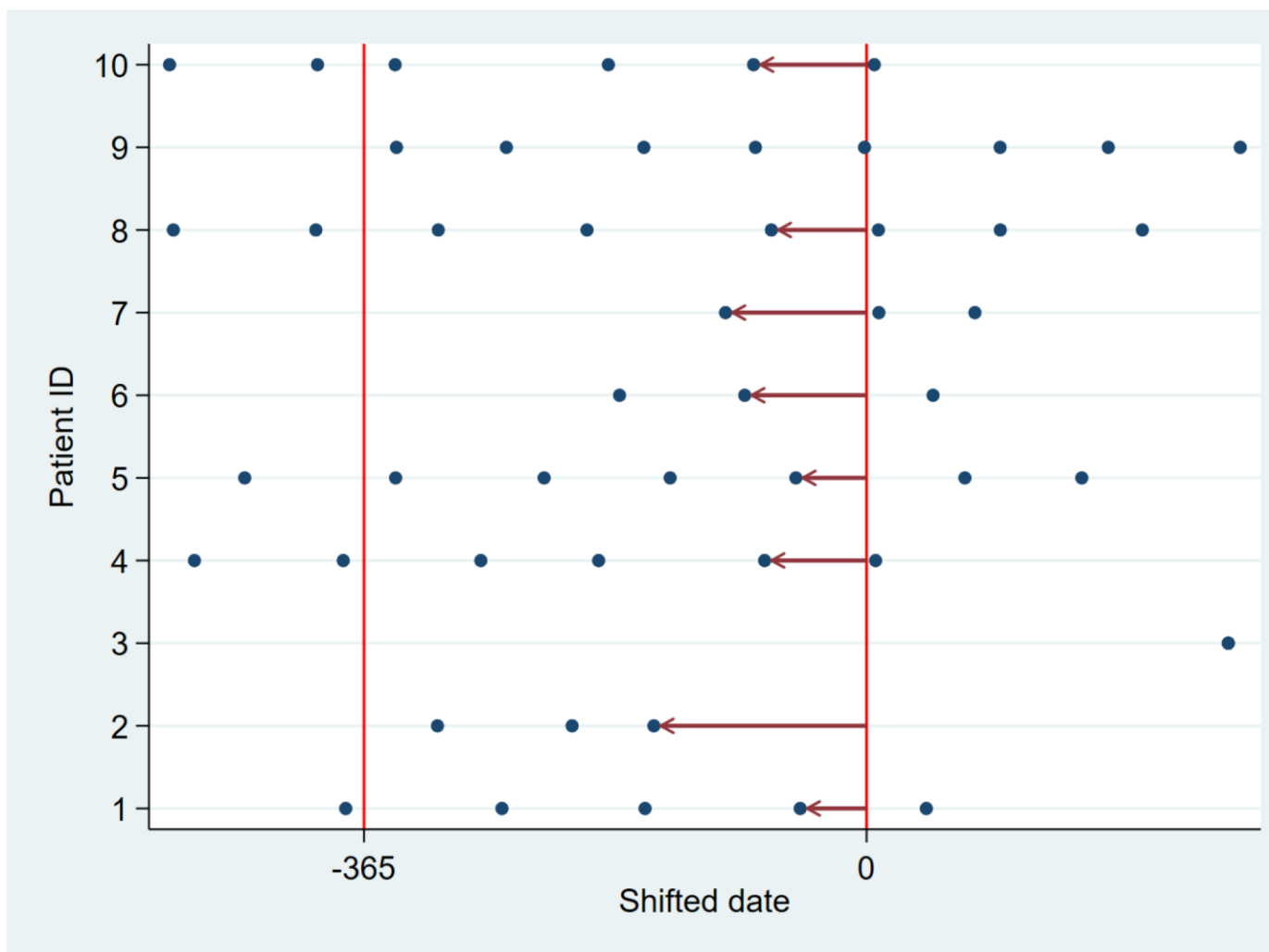
Idea: individual random index dates

- Step II: Find time since last prescription date



Idea: individual random index dates

- Step III – re-aligning index dates



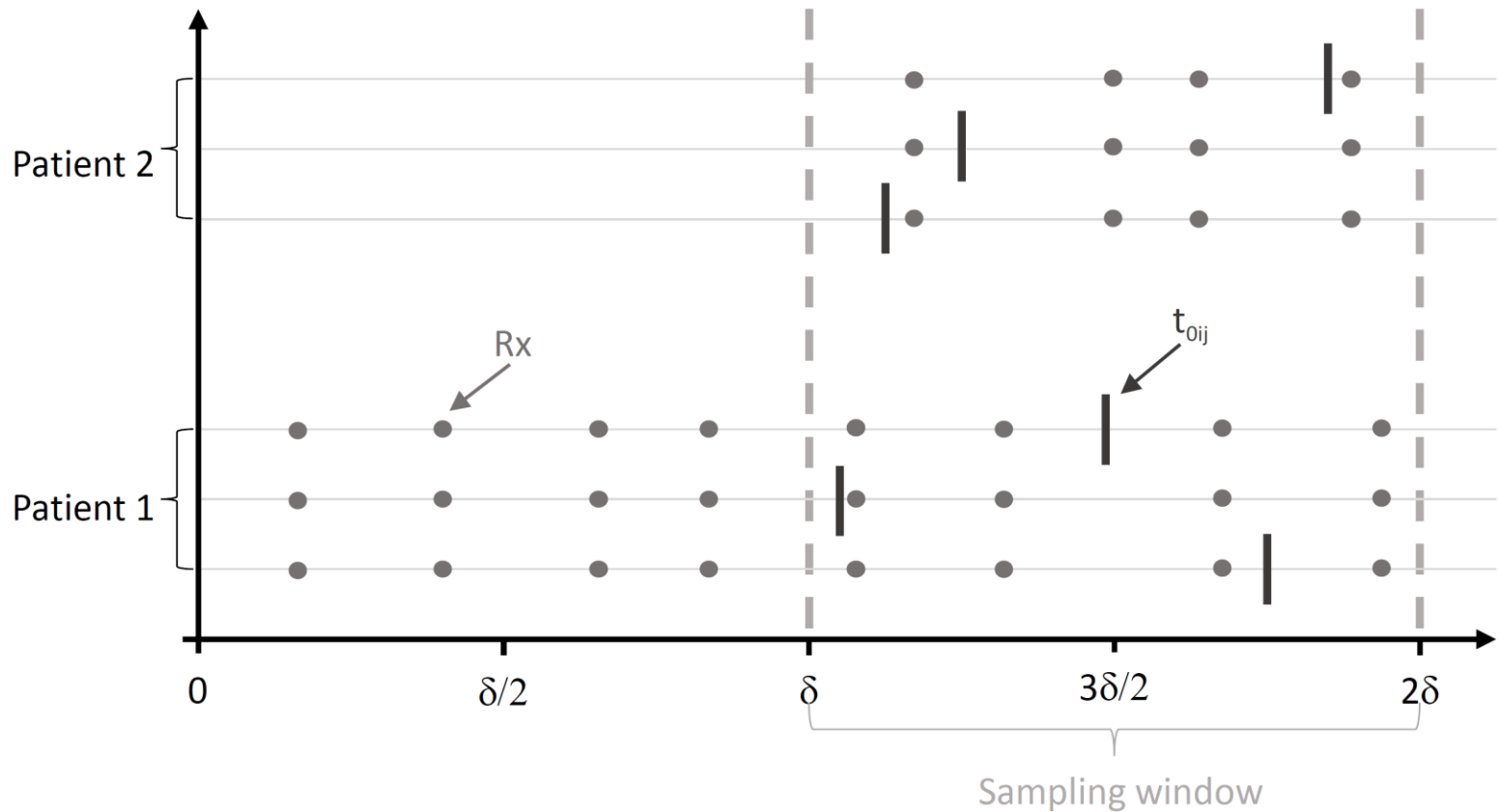
Empirical study revisited – Norway, warfarin

- A marginal estimate across the year
- Ordinary WTD and reverse WTD yield similar results

n	WTD	Index date	Estimated 80% percentile (days)
69,207	ordinary	Jan 1 2015	161.1 (159.7; 162.5)
77,674	reverse	Jan 1 2015	100.8 (99.7; 101.9)
79,850	reverse	Nov 1 2014	121.7 (120.6; 122.9)
82,892	reverse	Random	129.5 (128.3; 130.8)
73,836	ordinary	Random	133.6 (132.4; 134.8)

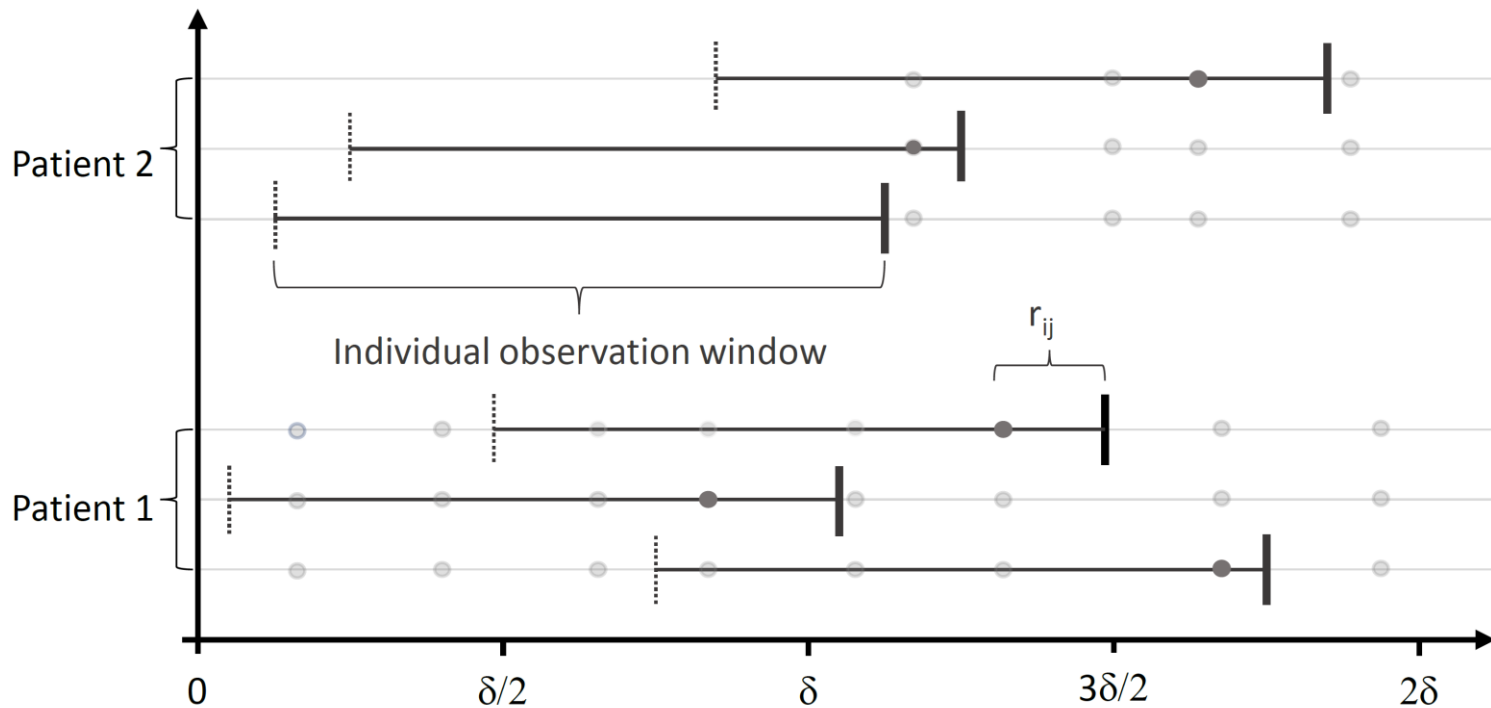
Multiple random index dates

- Can we improve statistical efficiency?
- Idea: Multiple random index dates for each individual



Multiple random index dates

- Robust variance estimates to handle dependency of observations from the same individual



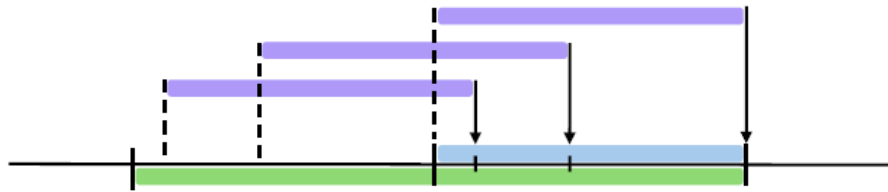
Multiple random index dates

- Shows higher efficiency in simulation studies, similar bias, lower coverage
- Improves precision in regression analyses

			1 random index date			5 random index dates			10 random index dates			20 random index dates		
Sex	Age (years)	n*	n _{pres} **	Estimated 80 th percentile in days (95% CI***)	Width CI (days)	n _{pres}	Estimated 80 th percentile in days (95% CI)	Width CI (days)	n _{pres}	Estimated 80 th percentile in days (95% CI)	Width CI (days)	n _{pres}	Estimated 80 th percentile in days (95% CI)	Width CI (days)
Female	0-49	843	635	75.8 (67.9; 83.8)	15.9	3165	74.4 (69.0; 79.8)	10.8	6341	76.8 (72.2; 81.5)	9.3	12662	77.2 (73.1; 81.4)	8.3
Female	50-59	641	520	79.9 (70.3; 89.6)	19.3	2559	79.4 (74.2; 84.6)	10.4	5114	80.0 (75.3; 84.6)	9.3	10205	79.9 (75.7; 84.2)	8.5
Female	60-69	1617	1315	83.6 (77.9; 89.3)	11.4	6619	84.9 (81.6; 88.2)	6.6	13252	86.4 (83.4; 89.3)	5.9	26562	86.1 (83.4; 88.7)	5.3
Female	70-79	3228	2763	96.5 (92.1; 101.0)	9.0	13880	96.2 (93.7; 98.8)	5.1	27720	96.7 (94.5; 98.9)	4.4	55428	96.4 (94.4; 98.3)	4.0
Female	80-89	3384	2887	111.3 (105.8; 116.8)	11.0	14429	106.9 (104.0; 109.9)	5.9	28911	107.0 (104.5; 109.5)	5.0	57717	107.0 (104.7; 109.2)	4.5
Female	90+	854	693	124.2 (110.1; 138.3)	28.1	3463	124.0 (116.5; 131.5)	15.0	6934	121.6 (115.2; 128.1)	12.8	13908	118.0 (112.5; 123.5)	11.0
Female	All	10567	8785	98.6 (95.9; 101.3)	5.4	44116	97.9 (96.4; 99.5)	3.1	88204	97.5 (96.1; 98.8)	2.7	176388	97.4 (96.2; 98.7)	2.5

4. Saturated sampling: all index dates

Saturated sampling



↓ Index date

- Look-back window: time window of pre-specified length within which we observe the time from the index date* to the last dispensation
- Sampling window: time window of pre-specified length within which we identify the index date(s)*
- Observation window: time window twice the length of the sampling window

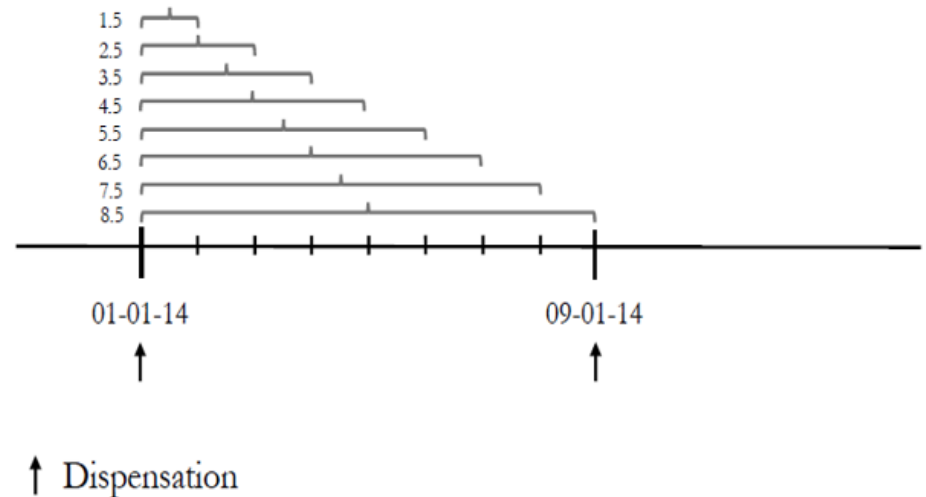
* Defined in different ways: i) single and fixed, ii) single and random, iii) multiple and random, iv) all days within the window (saturated version)

Saturated sampling

– expanding and collapsing the data

id	date	time span
1	01-01-2014	-
1	09-01-2014	8

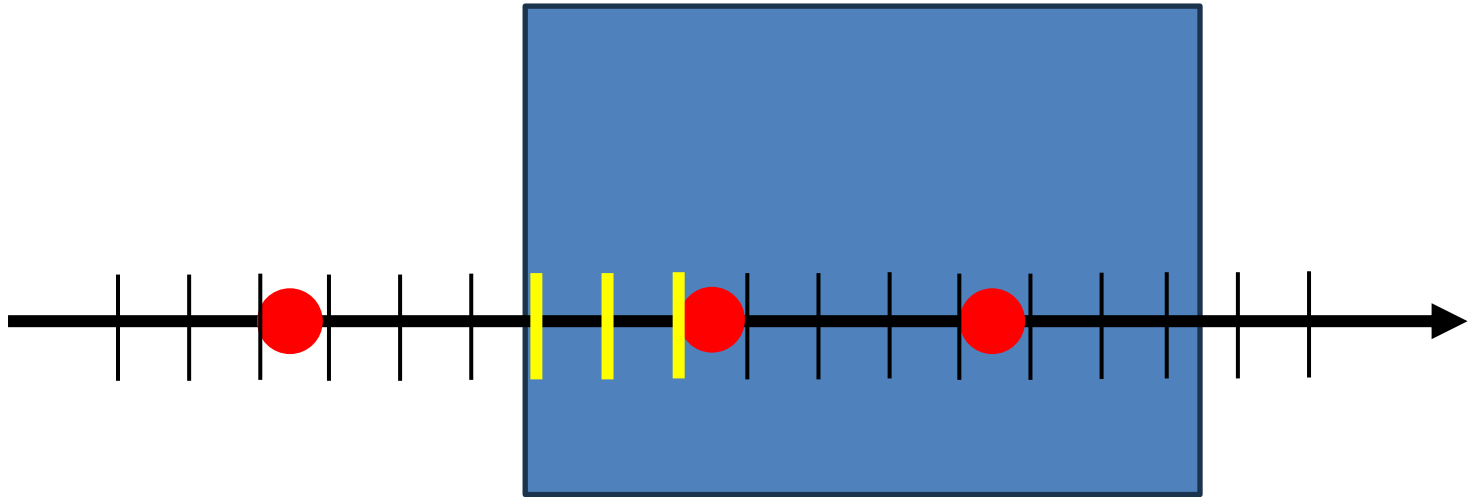
id	date	time
1	09-01-2014	0.5
1	09-01-2014	1.5
1	09-01-2014	2.5
1	09-01-2014	3.5
1	09-01-2014	4.5
1	09-01-2014	5.5
1	09-01-2014	6.5
1	09-01-2014	7.5



Saturated sampling

– expanding and collapsing the data

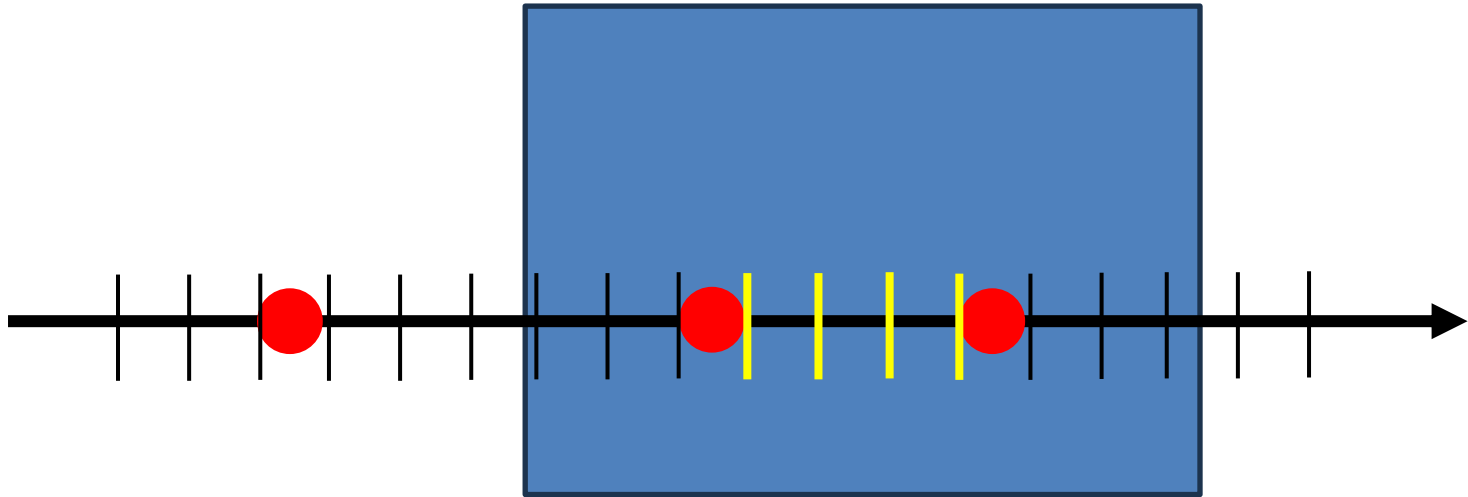
- Until first $R_x: \{4, 5, 6\}$



Saturated sampling

– expanding and collapsing the data

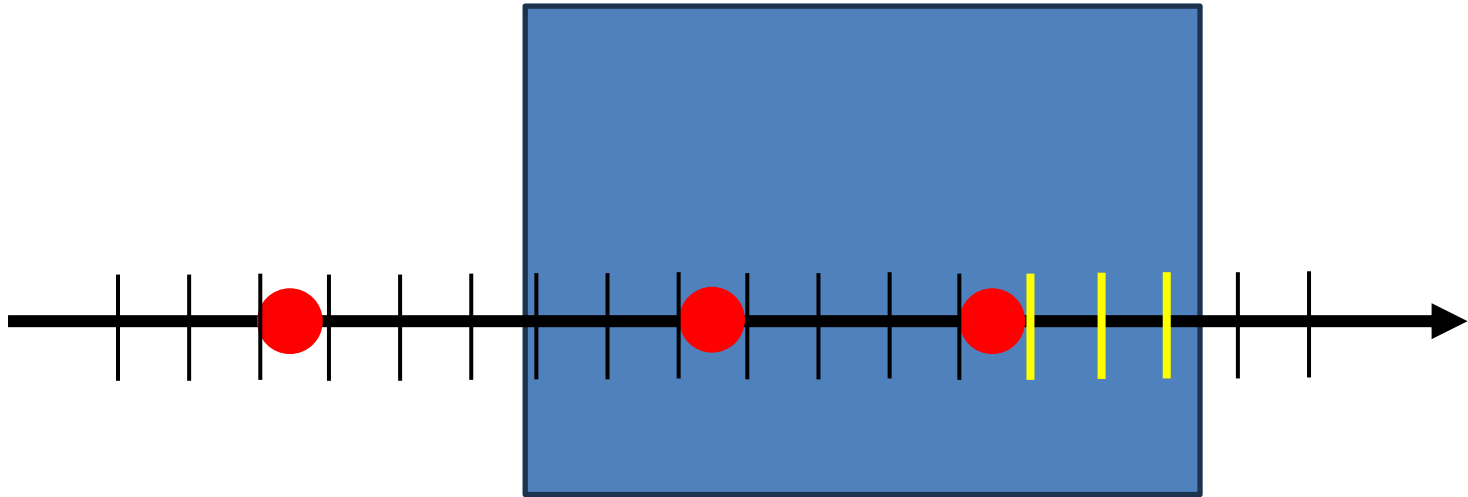
- Between first and second dispensing: {1,2,3,4}



Saturated sampling

– expanding and collapsing the data

- After last dispensing: {1,2,3}



Saturated sampling – counting

- Let $\{G_j\}_{j=0,\dots,m}$ be the set of gap times
 - G_0 is time from last dispensing in previous year
 - $G_1, \dots, G_{m-1}$ are gap times fully observed in sampling window
 - G_m is the last (censored) gap time: last dispensing to end of interval
- Assume largest gap time is $G_K = k$
- Assume that w_0 is distance from last dispensing before sampling window to first sampling date
- Then data set to be analyzed can be represented as

$$n_{0.5} = k - I(w_0 > 0.5) - \sum_{j=0}^m I(G_j < 0.5)$$

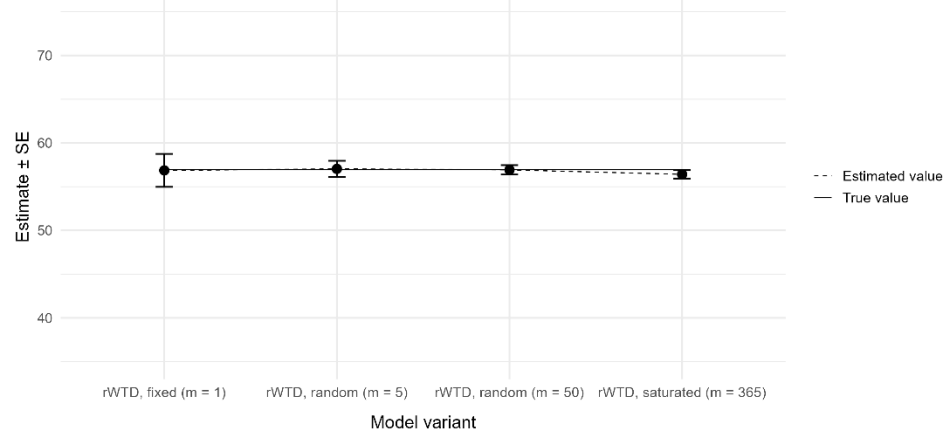
Saturated sampling – counting

- Let $\{G_j\}_{j=0,\dots,m}$ be the set of gap times
 - G_0 is time from last dispensing in previous year
 - $G_1, \dots, G_{m-1}$ are gap times fully observed in sampling window
 - G_m is the last (censored) gap time: last dispensing to end of interval
- Assume largest gap time is $G_K = k$
- Assume that w_0 is distance from last dispensing before sampling window to first sampling date
- Then data set to be analyzed can be represented as (and so on)

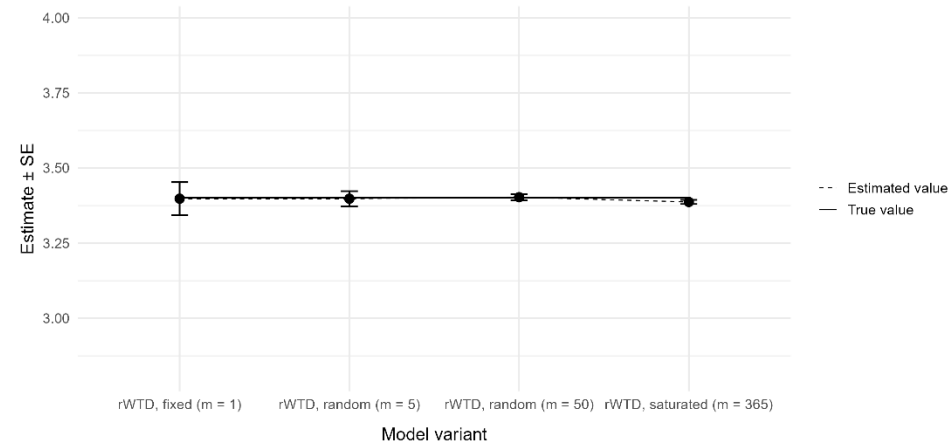
$$n_{1.5} = k - I(w_0 > 1.5) - \sum_{j=0}^m I(G_j < 1.5)$$

Saturated sampling – bias and precision

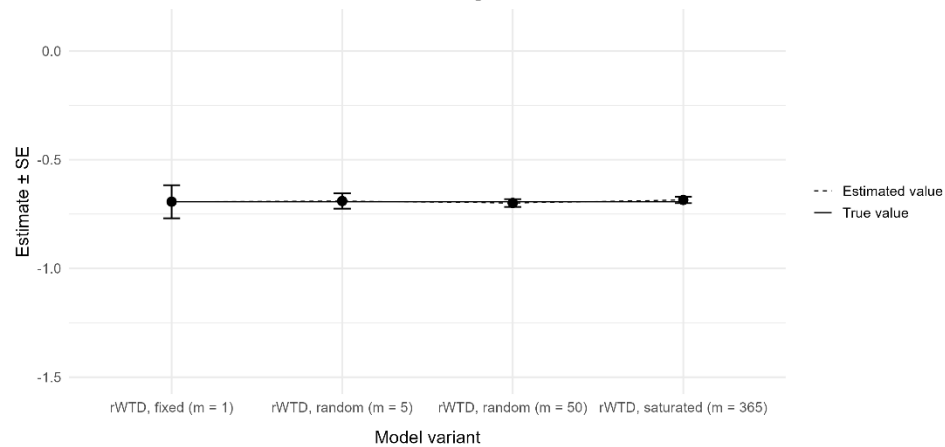
Dispensation lengths computed as 90th percentile of the IAD



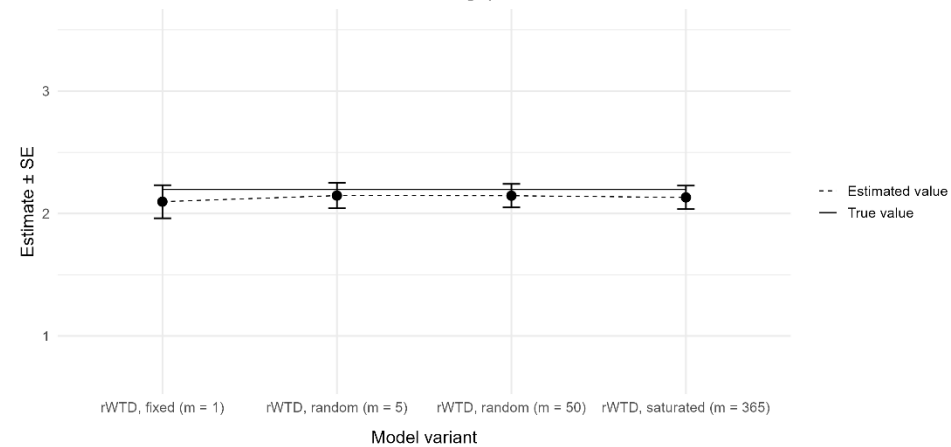
Estimated mu



Estimated Insigma



Estimated logitp



Conclusion

- **Main findings**

- The WTD approach provides a formal model for estimating prescription durations
- Flexible basis for expanding the model
- Possible to improve efficiency – but saturated sampling likely not worth the effort

- **Limitations**

- Vulnerable to misspecification
- Only works for usage patterns with continued treatment (multiple prescription redemptions)
- Lack of models integrated with exposure-outcome associations

A new perspective – where do we go?



“The Bean” (Millenium Cloud Gate), Chicago

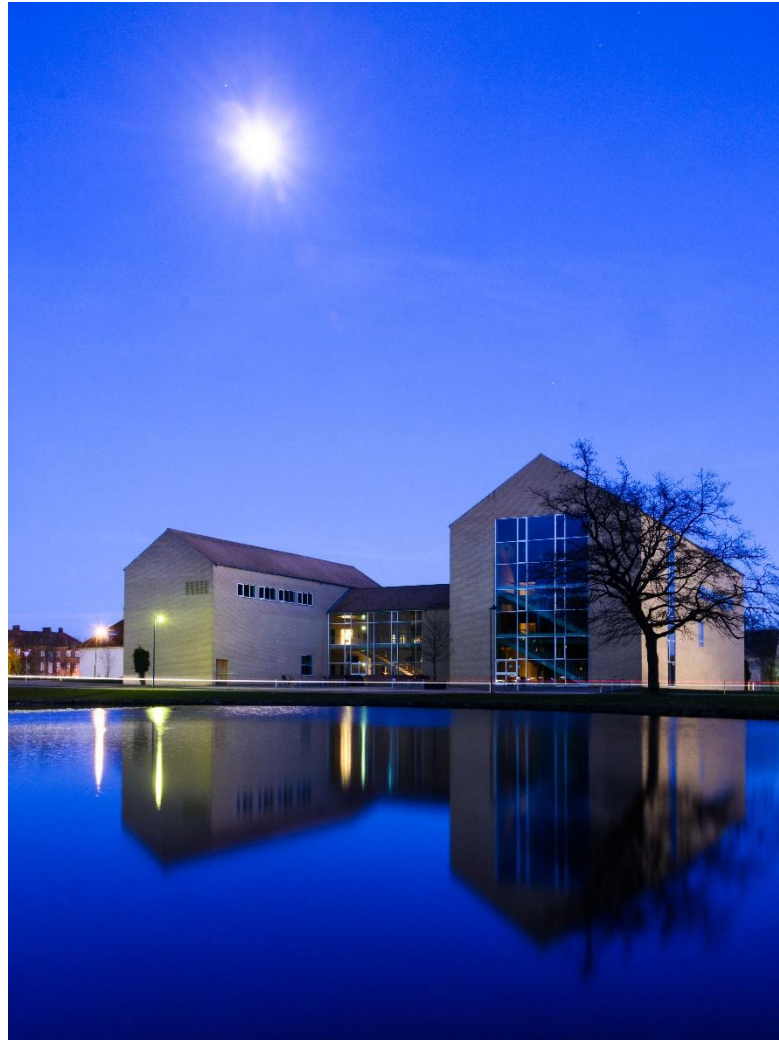
References

- Støvring H, Pottegård A, Hallas J. Determining prescription durations based on the parametric waiting time distribution. *Pharmacoepidemiol Drug Saf.* 2016;25(12).
- Støvring H, Pottegård A, Hallas J. Estimating medication stopping fraction and real-time prevalence of drug use in pharmaco-epidemiologic databases. An application of the reverse waiting time distribution. *Pharmacoepidemiol Drug Saf.* 2017;26(8):909–16.
- Støvring H, Pottegård A, Hallas J. Refining estimates of prescription durations by using observed covariates in pharmacoepidemiological databases: an application of the reverse waiting time distribution. *Pharmacoepidemiol Drug Saf.* 2017;26(8).
- Hallas J, Pottegård A, Støvring H. Using probability of drug use as independent variable in a register-based pharmacoepidemiological cause-effect study—An application of the reverse waiting time distribution. *Pharmacoepidemiol Drug Saf.* 2017;26(12):1520–6.
- Thrane JM, Støvring H, Hellfritzsche M, Hallas J, Pottegård A. Empirical validation of the reverse parametric waiting time distribution and standard methods to estimate prescription durations for warfarin. *Pharmacoepidemiol Drug Saf.* 2018 Jun 28;

References (continued)

- Bødkergaard, K. *et al.* Using the waiting time distribution with random index dates to estimate prescription durations in the presence of seasonal stockpiling. *Pharmacoepidemiol. Drug Saf.* **29**, 1072–1078 (2020).
- Støvring H, Pottegård A, Hallas J. A new likelihood model for analyses of pharmacoepidemiologic case-control studies which avoids decision rules for determining latent exposure status. *BMC Medical Research Methodology* (Accepted, 2021)

Thanks for your attention – questions welcome!



(Aarhus University, March 2016 – H Støvring)

A1. Validation study

Validation (Thrane, 2018)

- Thrombobase:
Records information of patients treated with anticoagulants in Southern Denmark
- Patient characteristics
 - have suffered a thrombosis
 - have their INR monitored
 - discard the first 4 months of treatment (stabilized regime)
 - -> high compliance with treatment guidelines
- Included 2,848 patients treated between 1998 and 2010
- Records prescribed dose for each prescription
-> we considered this *THE Gold Standard*

*Thrane JM, Støvring H, Hellfritsch M, Hallas J, Pottegård A. Empirical validation of the reverse parametric waiting time distribution and standard methods to estimate prescription durations for warfarin. *Pharmacoepidemiol Drug Saf.* 2018 Sep;27(9):1011–8.

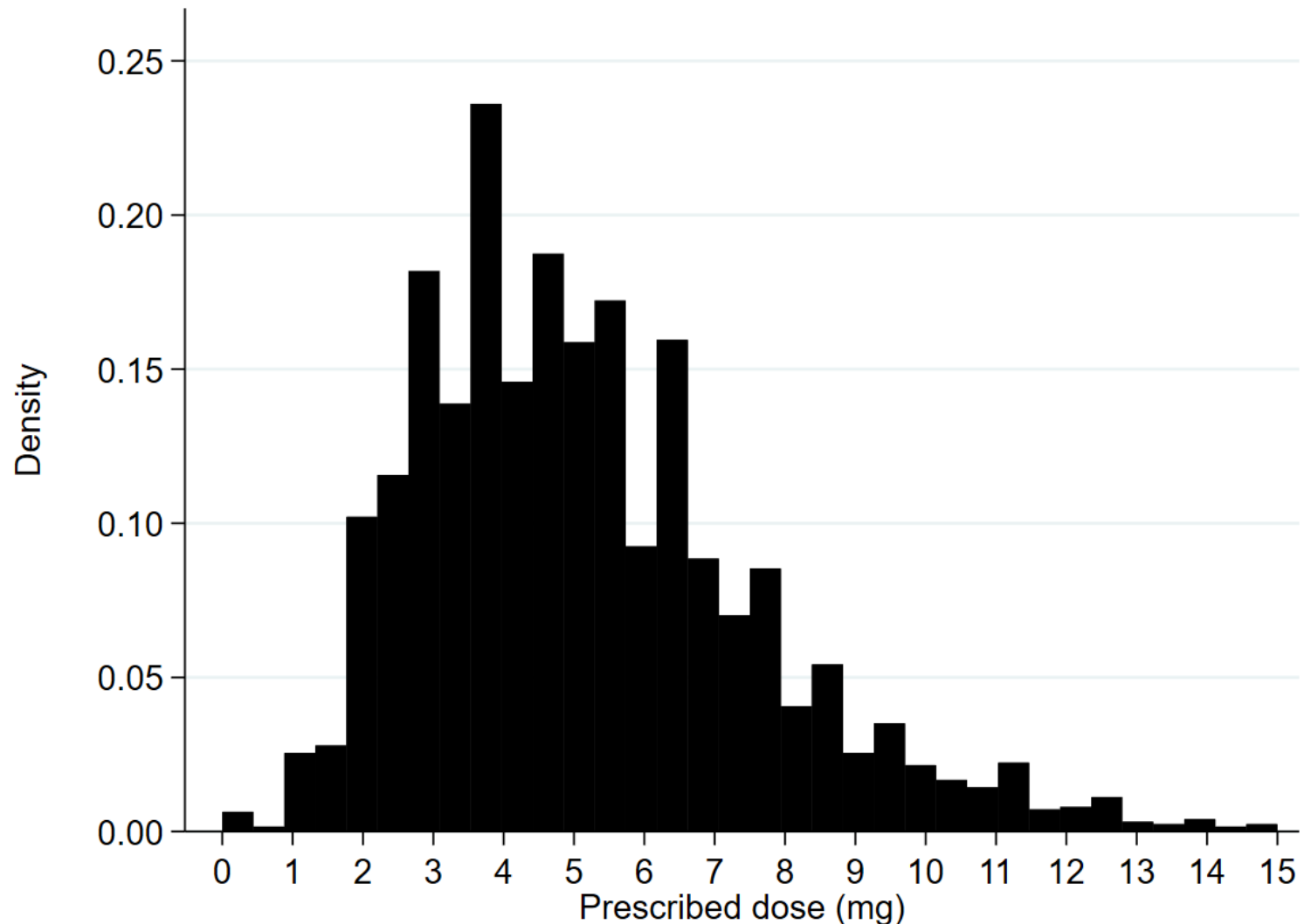
Validation study - design

- Methodologies to be validated
 - 1 DDD per day
 - 0.5 DDD per day
 - rWTD estimated mean duration with or without covariates (sex, age, # packages, time to last Rx)
- Computed estimated dosage for rWTD as

$$\frac{\text{total dose obtained in a prescription redemption}}{\text{estimated duration}}$$

- Compare with prescribed dosage *on log-scale*
- We report bias and ratio of limits-of-agreement (95% PI of deviations)

Distribution of prescribed daily doses *(considered to be the Gold standard)*



Validation results

Method	Covariates in model?	Relative bias	Limits of Agreements	Limits of Agreements Ratio
1.0 DDD	N/A	62%	-38% to 319%	6.721
0.5 DDD	N/A	-19%	-69% to 110%	6.721
rWTD (mean)	None	-0%	-62% to 162%	6.857
	Sex, age	1%	-60% to 155%	6.382
	Sex, age, #packs	2%	-57% to 146%	5.778
	Sex, age, #packs, time since last Rx	2%	-49% to 104%	4.036

- Reverse WTD is virtually unbiased
- Precision increases with more covariates
- Low precision of individual predictions, even with covariates

