Bayesian Drug Disease Model with Stan

Using published longitudinal data summaries in population models

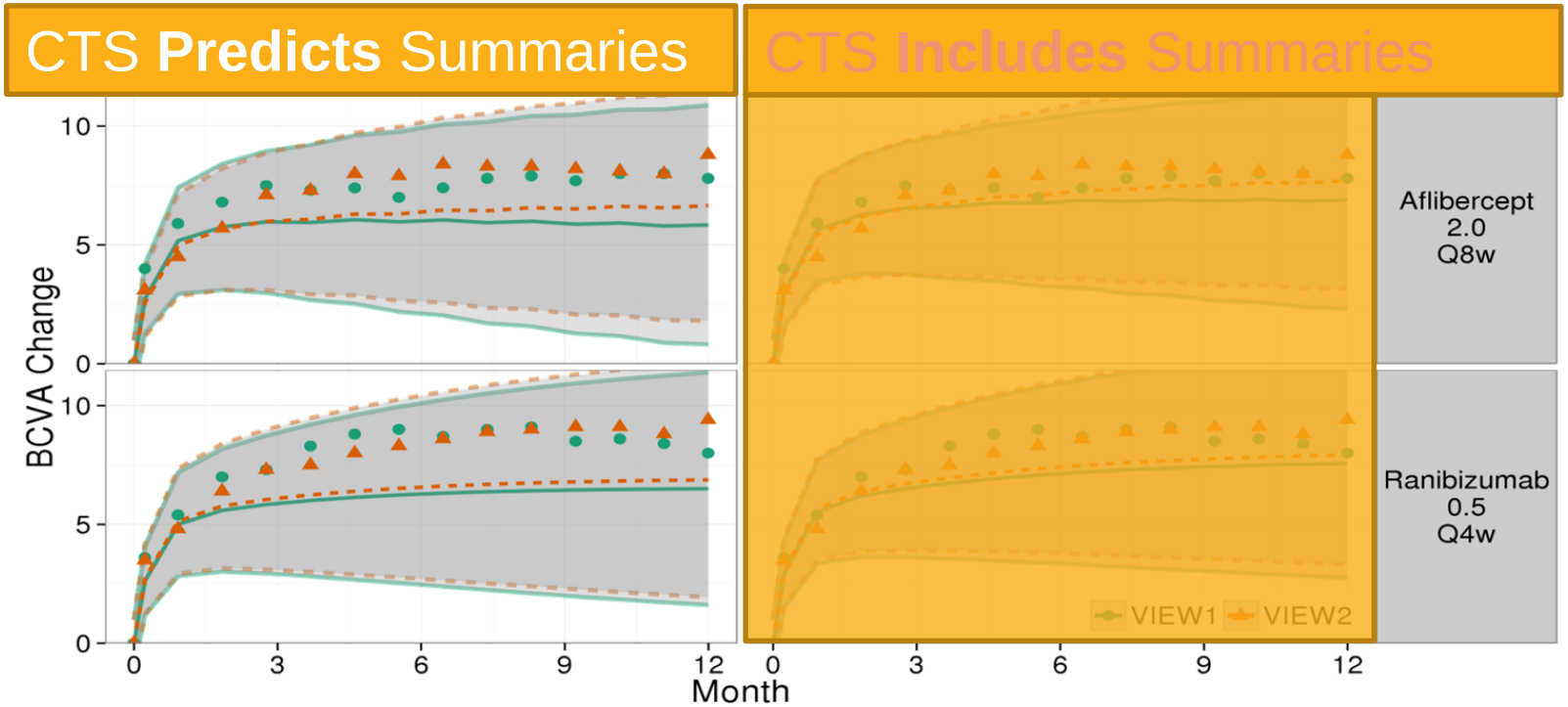
S. Weber[1], B. Carpenter[2], D. Lee[2], F. Y. Bois[3], A. Gelman[2], A. Racine[1]

(1) Novartis, Basel; (2) Columbia Uni., New York; (3) Uni. de Technologie de Compiègne

23. PAGE meeting, 11. June 2014, Alicante, Spain

NOVARTIS

Data Summary from Source VIEW 1+2, Heuer et al, Opht. Vol. 119, 2012



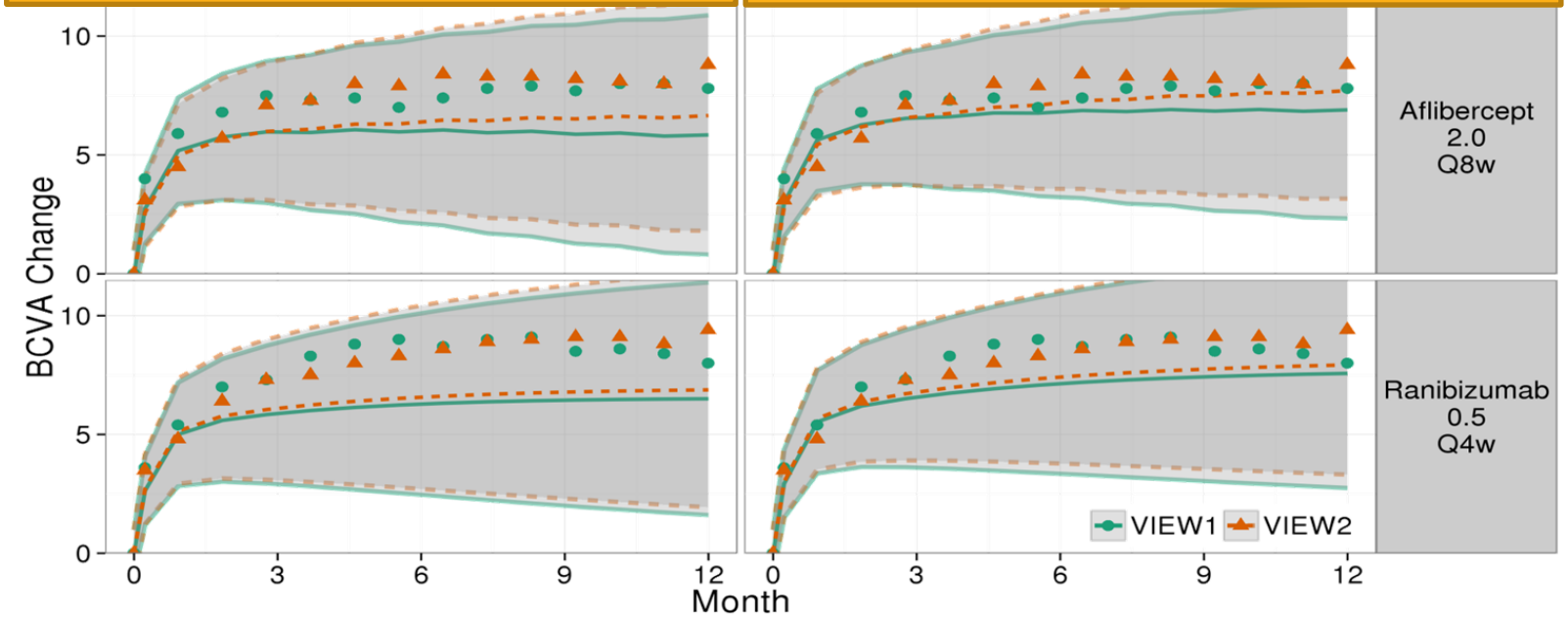
Trial Design using Clinical Trial Simulation (CTS)
From Patient Data (left)

NOVARTIS

Data Summary from Source VIEW 1+2, Heuer et al, Opht. Vol. 119, 2012

CTS **Predicts** Summaries

CTS **Includes** Summaries



Trial Design using Clinical Trial Simulation (CTS)
From Patient Data (left) ***and*** *Data Summaries (right)*

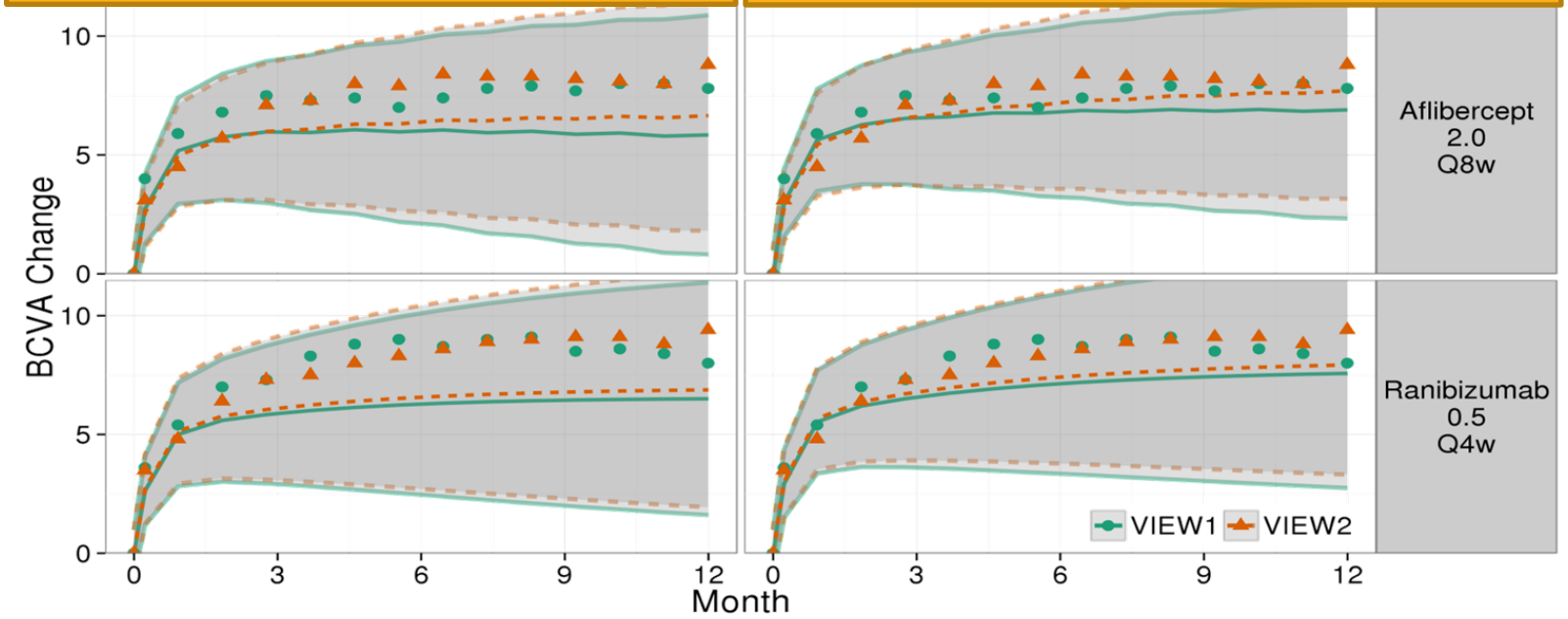
NOVARTIS

Outline

Data Summary from Source VIEW 1+2, Heuer et al, Opht. Vol. 119, 2012

CTS **Predicts** Summaries

CTS **Includes** Summaries



Trial Design using Clinical Trial Simulation (CTS)
From Patient Data (left) ***and*** *Data Summaries (right)*

1. Bayesian Drug Disease Model with Stan
2. Learning From Published Data Summaries with Sample Importance Resampling (SIR)

NOVARTIS

Example: Wet Age Macular Degeneration (AMD)

If Untreated Severe Degradation of Vision within ~1-2y

- Prevalent in elderly patients
- Treatment with intravitreal injections of anti-VEGF inhibitor
  - Ranibizumab
  - Aflibercept
- Clinically relevant endpoint is change from baseline BCVA Δ

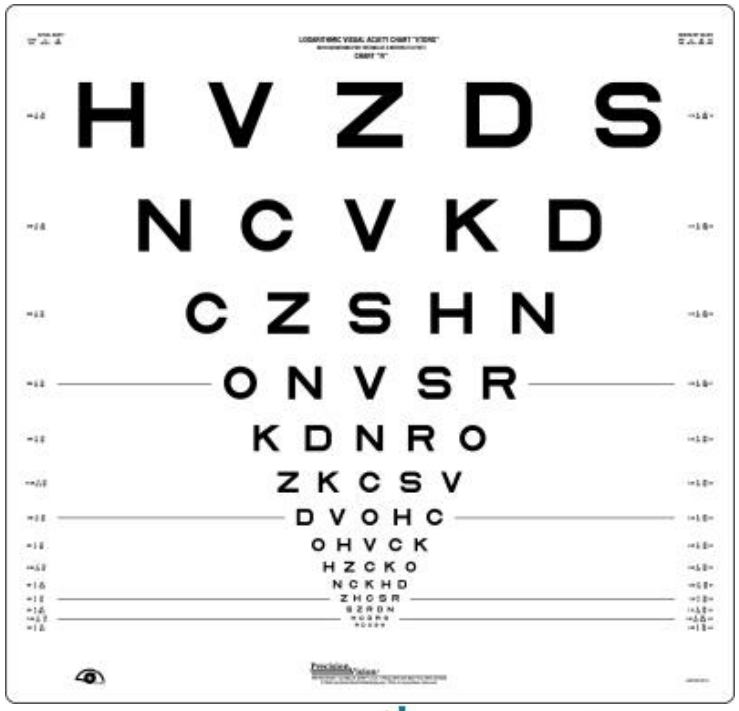
BCVA
best corrected visual acuity measured as letters read from ETDRS chart

Decrease

Linearly

on

Log-Scale



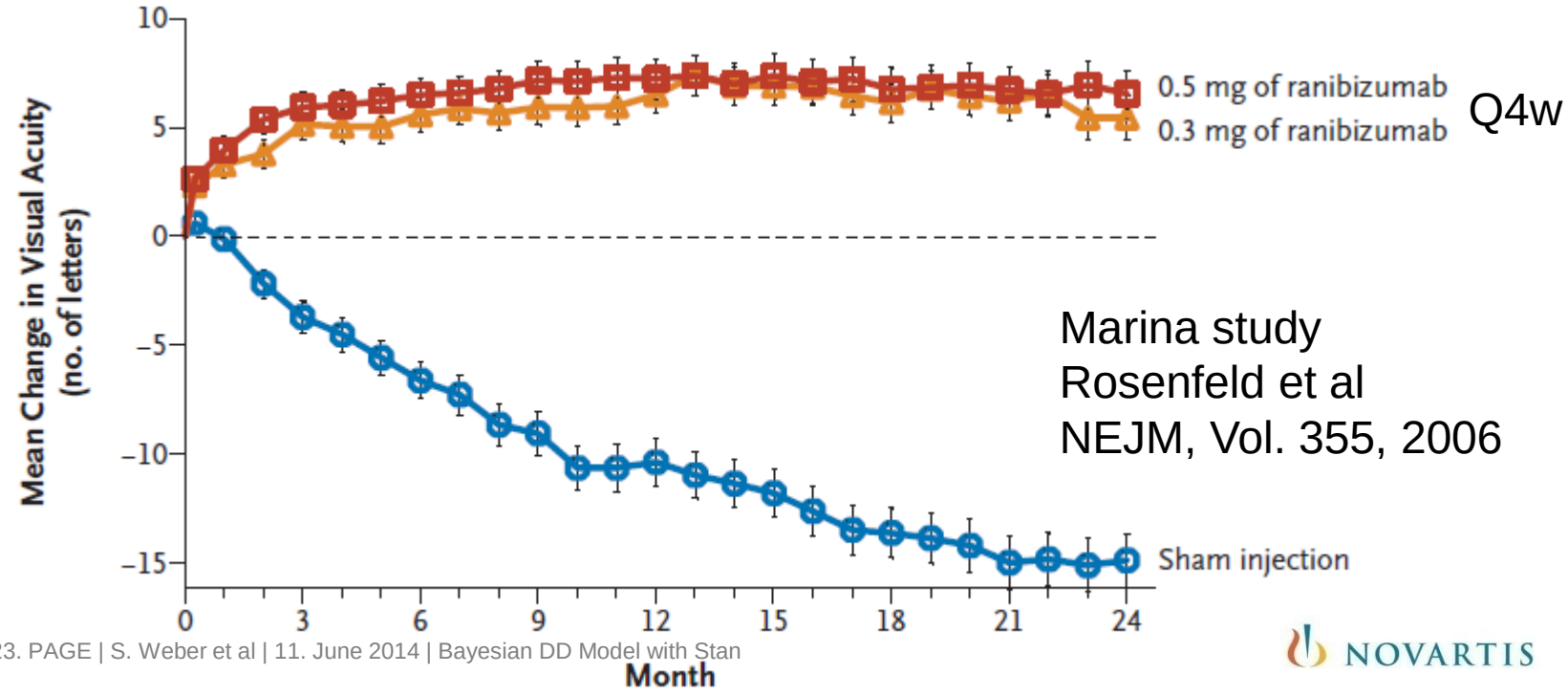
(C) Precision Vision

NOVARTIS

Available Patient Data on Ranibizumab

In-House Conducted Studies Marina, Anchor and Excite

- Patient Level Data
  - ~ 200 Placebo
  - ~1100 Ranibizumab
- Baseline
  - BCVA 53 (± 13) letters
  - Age 73 (± 7.5) years
- Dose 0.3/0.5mg Q4/12w
- BCVA every 4w



NOVARTIS

Drug Disease Model For BCVA Letter Count

A Turnover K-PD Model with Stimulation of k_{in}

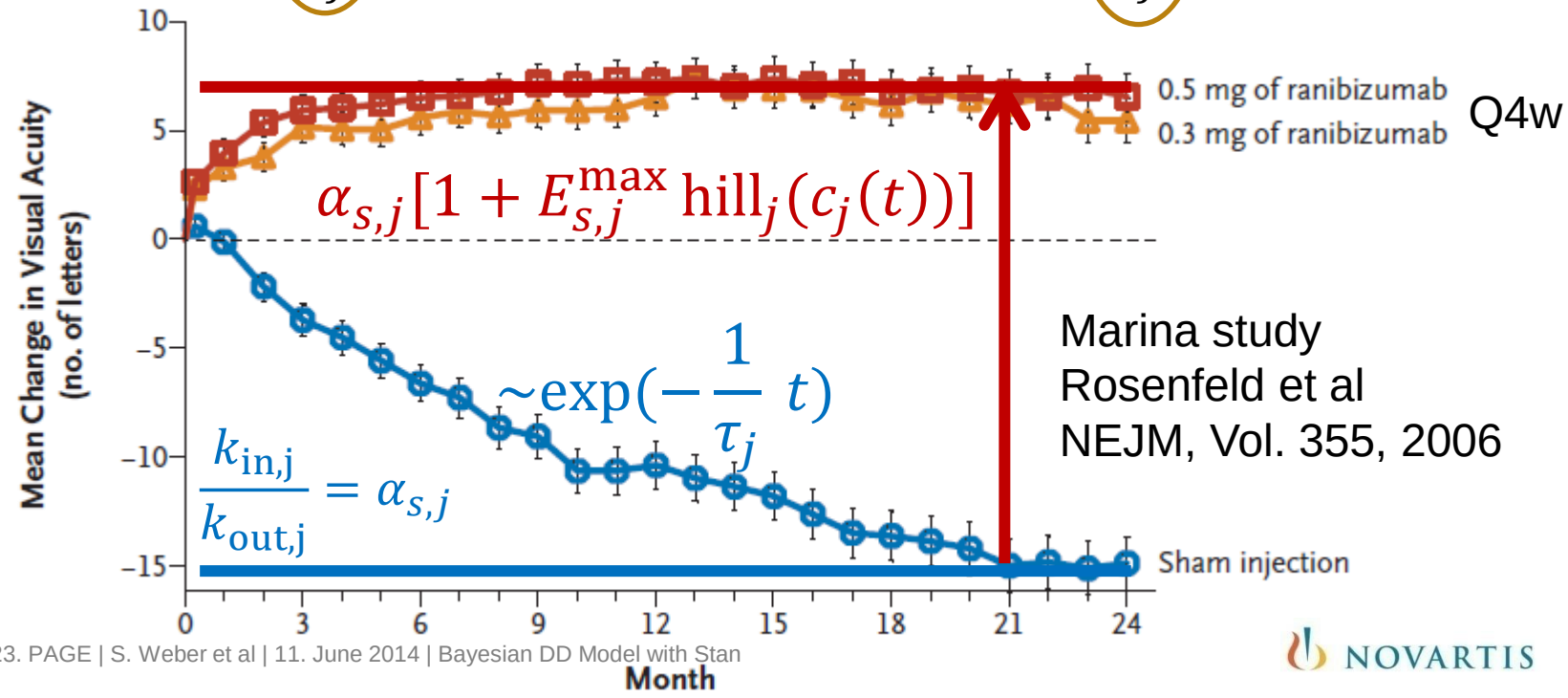
Ocular PK absent
V := 4mL
$T_{½}$:= 9d

$$\text{loglog}_{s,a}(y_{j,i}/100) \sim \text{Normal}(g_j(t_{j,i}), \sigma_y)$$

BCVA letter count $0 \le y_{j,i} \le 100$ for patient j at time point $t_{j,i}$

$k_{in,j}$ $k_{out,j}$

$$\frac{dg_j(t)}{dt} = \frac{\alpha_{s,j}}{\tau_j}\left[1 + E_j^{\max}(t)\,\text{hill}_j(c_j(t))\right] - \frac{1}{\tau_j}\,g_j(t)$$



Q4w

Marina study
Rosenfeld et al
NEJM, Vol. 355, 2006

NOVARTIS

Key Modeling Choices

- Subject-specific random effects
  - Baseline BCVA $g_j(0) = \alpha_0 + \eta_j^0$
  - Placebo steady-state BCVA $\frac{k_{\text{in,j}}}{k_{\text{out,j}}} = \alpha_{s,j} = \alpha_s + \eta_j^s$
- Efficacy varied substantially between and within a study → **study/treatment arm random effect $\boldsymbol{\eta_{k(j)}^{sa}}$**

$$\log\left(E_{s,j}^{\max}\right) = lE_s^{\max} + \{covariates\} + \eta_{k(j)}^{\text{sa}}$$

- **Informative priors for stable model fit**
  - **Inclusion/exclusion criteria: Baseline BCVA α_0 = 37-75 letters**
  - **Clinical disease knowledge: Loss of 14-16 letters + τ ~ ½ – 1½y**

Model Inference with Stan to Obtain Posterior

Stan Extended for ODE Support in Cross-Divisional Project at Novartis

Patient Data x

Priors $p(\theta)$

«Turnover» ODE-Stan 2.2

K-PD Turnover Model Likelihood $p(x \mid \theta)$

$\{\theta_1, \ldots, \theta_{200}\}$

$\{\theta_{201}, \ldots, \theta_{400}\}$

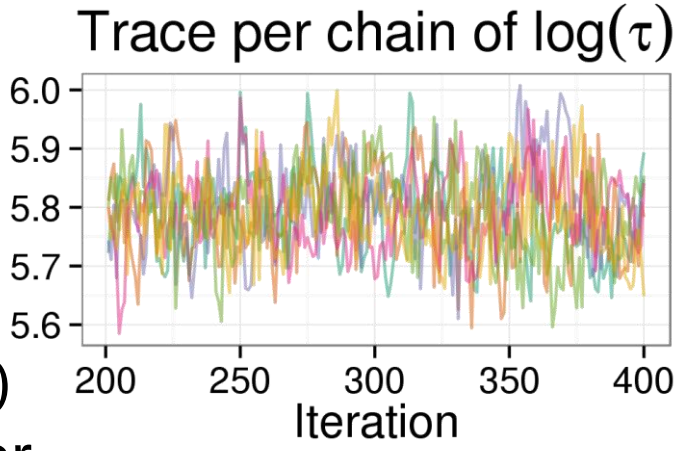
$\{\theta_{401}, \ldots, \theta_{600}\}$

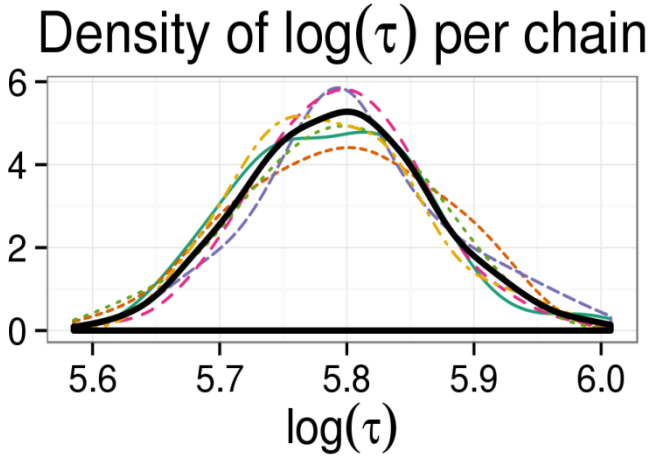
$\{\theta_{601}, \ldots, \theta_{800}\}$

$\{\theta_{801}, \ldots, \theta_{1000}\}$

Run 5 chains in parallel ~22-30h

$p(\theta \mid x)$ Posterior





$R_{hat} < 1.02$

$230 < N_{eff} < 1000$

http://mc-stan.org/

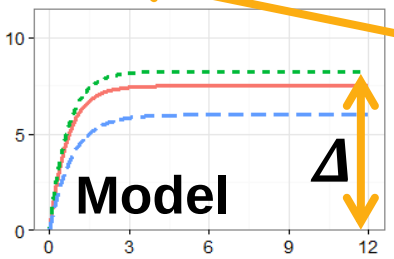
NOVARTIS

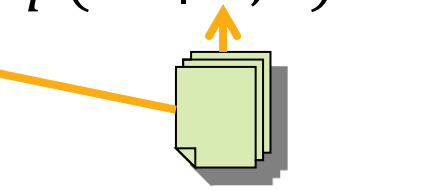
Learning From Published Data Summaries

Expand Model to Further Substances from Published Data Summaries

- Substances differ by EC50 (and maybe Emax)
- Strategy: Bridge available knowledge to published data summaries and learn by applying
  Sample Importance Resampling (SIR) / «discrete» Bayes

Available Knowledge	**1. Discount**	**2. Bridge**	**3. Learn**
Patient Data x	Same Emax Vague EC50	Predict Mean BCVA Change Δ	Bayes' Theorem
$p(\theta \mid x)$	$p(\theta^\star \mid x)$	$p(\Delta \mid \theta^\star, x)$	$p(\theta^\star \mid \Delta, x)$





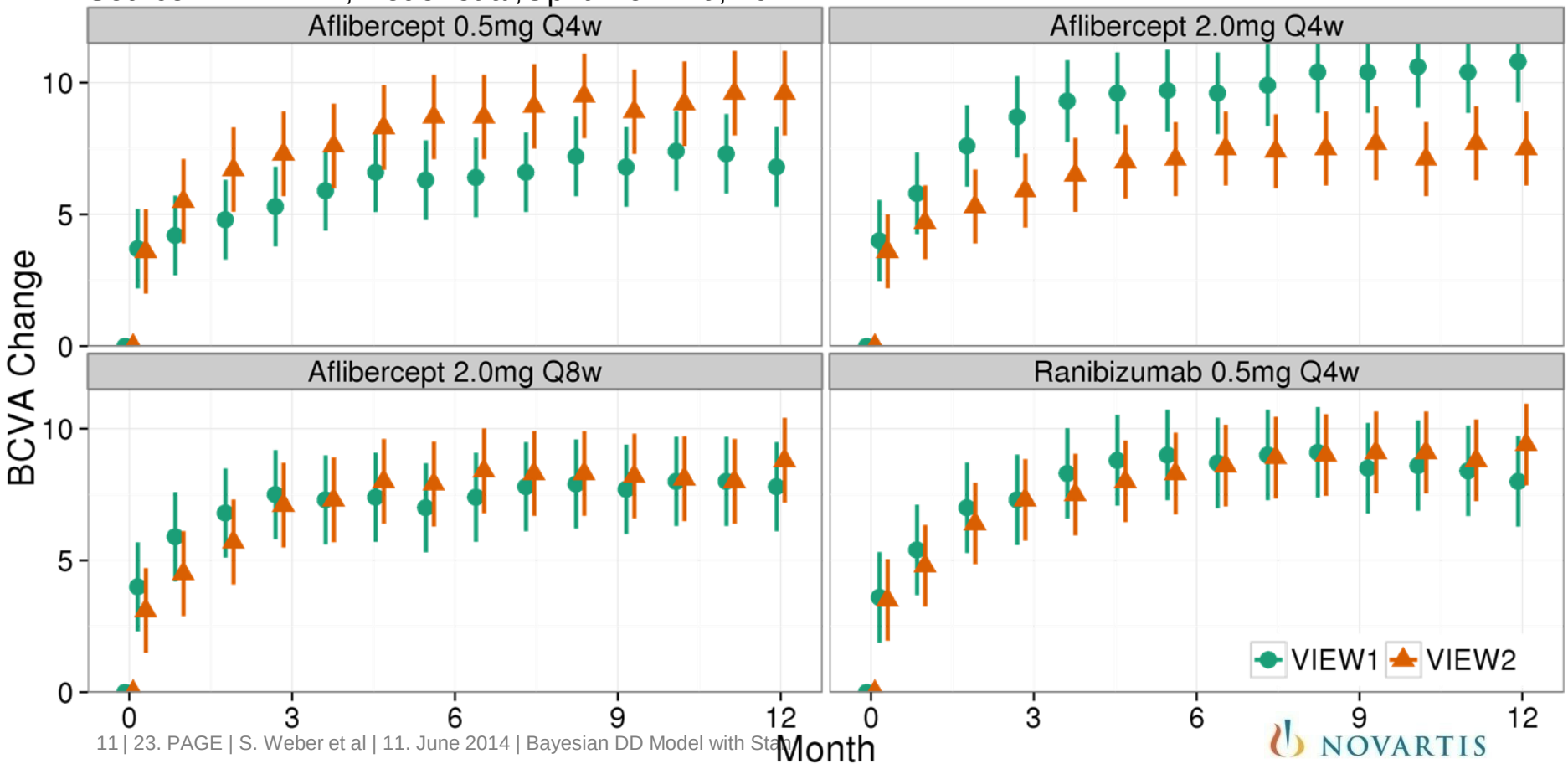
NOVARTIS

Available Summary Data on Marketed Substance

Aflibercept NI trial VIEW1+2 with Ranibizumab control

- Each treatment arm included ~300 patients → ~2400
- Large variation between (some) replicated treatment arms

Source: VIEW 1+2, Heuer et al,Opht. Vol. 119, 2012



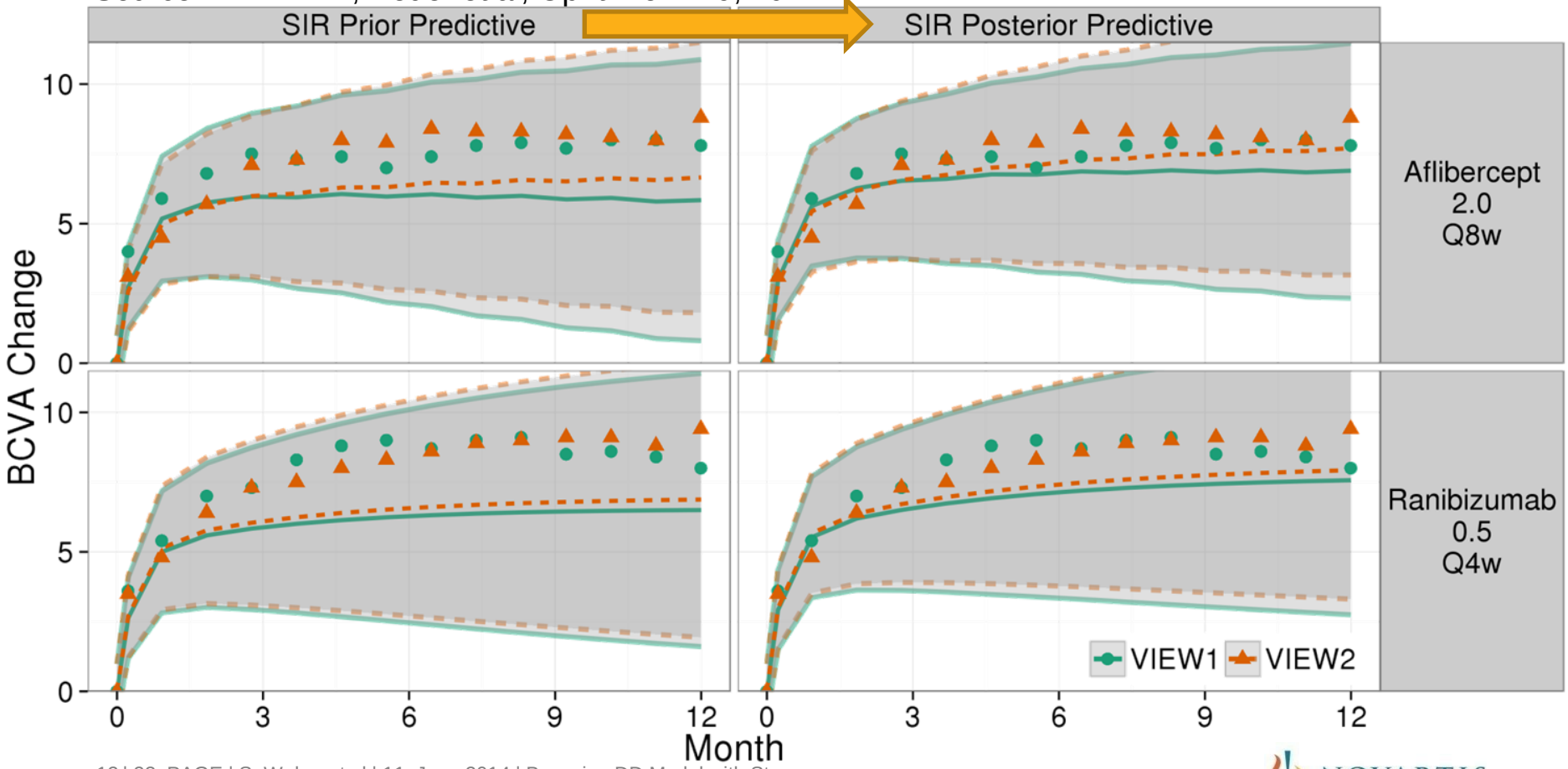
NOVARTIS

Learning From Published Data Summaries with SIR

Predicted Mean BCVA Change with 95% CI Before and After SIR

- **Updated Emax and EC50**
- **Precision of Emax increased by ~40% for Aflibercept**

Source: VIEW 1+2, Heuer et al, Opht. Vol. 119, 2012



NOVARTIS

Summary

Bayesian Drug-Disease Model Integrates Patient and Summary Data

- Bayesian drug-disease model predicts time-profiles of drug responses using patient and summary data.
  Valuable to plan non-inferiority trials with a *control* which uses a *substance for which only summary data are available.*

- Sparse sampling makes model(s) quickly hard to fit without prior information.
  ***Bayesian framework Stan* coupled with informative priors was key for a stable model fit.**

- ***Sample Importance Resampling*** allows to obtain **Emax and EC50 of further substance from data summaries;** performs a «discrete» application of Bayes' Theorem.

Acknowledgements

Colleagues at Novartis
Great Team

- Amy Racine
- David James
- Ramesh Sarangapani
- Beat Neuenschwander
- Simon Wandel
- Satrajit Roychoudhury

Stan Team
For creating ODE-Stan

- Andrew Gelman
- Bob Carpenter
- Daniel Lee
- Frederic Y. Bois
- Michael Betancourt

NOVARTIS

References

- Marina
  - Rosenfeld et al, NEJM, 2006, 355:1419
- VIEW1+2
  - Heier et al, Opht., 2012,119(12):2537
- Stan
  - Stan Dev Team 2014, Version 2.2. http://mc-stan.org.
  - Hoffman and Gelman, JMLR., 2014, Vol. 15, 1351
- SIR / Approximate Bayesian Computation
  - Rubin, Ann. Stat., 1984, Vol. 12, 1151
  - Smith and Gelfand, Am. Stat., 1992, Vol. 46, No. 2
  - Diggle and Gratton, J Roy. Stat. Soc., 1984, Series B 46:193

Extra Material

NOVARTIS

Sample Importance Resampling

1. Extend draw i from posterior $\theta_i^\star = (\theta_i, lE_{s,R}^{\max,\star}, l\mathrm{EC50}_A^\star, lE_{s,A}^{\max,\star})$
2. Simulate expected profiles of ***n*** patients per treatment arm of study VIEW1+2 with draw i and replicate **m** times $\Delta_{k,j,i} = \mathrm{loglog}_{s,a}^{-1}(g_{j,k(j),i}(1y)) - \mathrm{loglog}_{s,a}^{-1}(g_{j,k(j),i}(0))$
3. Calculate baseline change at 1y; estimate between *patient **j*** and between *study-treatment arm **k*** variance $\mathrm{E}(1/\omega_{\mathrm{sa},i}^2)^{-1} = m^{-1}\sum_{k=1}^{m}(\Delta_{k,*,i} - \Delta_{*,*,i})^2$ $\mathrm{E}(1/\omega_{\Delta,i}^2)^{-1} = (m\,n)^{-1}\sum_{k=1}^{m}\sum_{j=1}^{n}(\Delta_{k,j,i} - \Delta_{k,*,i})^2$
4. Repeat step 1-3 for all draws from the posterior such that we obtain the posterior of the baseline change
5. Update the posterior via importance resampling, i.e. consider posterior from model as prior and update with summary level data from VIEW1+2. Weight given by **prior predictive** $p(\Delta_{*,*}^{\mathrm{VIEW}} \mid \theta_i^\star) = \mathrm{N}(\Delta_{*,*,i}, (\omega_{\Delta,i}^2 + \overline{\sigma}_{y,i}^2)/n + \omega_{\mathrm{sa},i}^2)$

NOVARTIS

Sample Importance Resampling

- Duality between sample and density
- Rejection sampling reweights samples, i.e. sample from density $g(\theta)$ samples with density $f(\theta)$
  - Let $M = \sup_\theta \frac{f(\theta)}{g(\theta)}$
  - Draw $\theta \sim g(\theta)$ and $u \sim (0,1)$
  - If $u \leq \frac{f(\theta)}{M\, g(\theta)}$, then accept θ; otherwise reject it
- Bayes Theorem $p(\theta \mid x) = \frac{p(x|\theta)\, p(\theta)}{p(x)} \propto p(x \mid \theta)\, p(\theta)$
  Equivalent to reweighting prior sample by likelihood
  - Generate prior sample with density $g(\theta) = p(\theta)$
  - Resample prior sample with $f_x(\theta) = p(x \mid \theta)\, p(\theta) \rightarrow \frac{f_x(\theta)}{p(\theta)} \propto p(\theta \mid x)$

SIR Assumptions & Limitations

- External data generated by the same
  - Structural model - only few parameters change
  - Statistical model - variance components are identical
- Approximate method
- Update based on deviations of predicted (model) mean from summary level data
- Only reliable under steady-state conditions

Stan Model Code Excerpt

- Stan models are declared with up to 6 blocks
  Data, transformed data, **parameters**, transformed parameters, **model**, generated quantities

- In transformed parameters IPRE is computed

```
for (j in 1:J) { // individual parameters are set here
 real g0;
 g0 <- alpha_0[SARM[j]] + eta_bva[j];
 if(PBO[j]) { // analytic solution for placebo (no drug)
    real asymp;
    asymp <- kin / kout;
    for(i in 1:N[j])
        Lypred[y_index + i - 1] <- asymp + (g0 - asymp) * exp(-
kout * t[y_index + i - 1]);
```

- The model block contains priors + sampling statements

```
ltau ~ normal(log(365.), log(1.5)/1.96);
// vectorized likelihood
Ly ~ normal(Lypred, sigma_y);
```