Cellular Impact of Contrast Agents

Angela Ariza de Schellenberger, Eric Barnhill and Shravan Vasishth

2018.05.14

Overview

This study evaluated whether novel cell measurement technologies showed differences in cell behavior between three widely used MRI contrast agents: Gadovist, Magnevist and Dotarem. These three agents are considered to have differing levels of safety and toxicity and further insight into the relative biological impact of these agents could have immediate impact for MRI protocols worldwide. To evalaute the relative toxicitiy of these agents we investigated whether we could detect differences between their effects on varying kinds of human cells. We measured differences with two different measurement techniques: Time-Of-Flight Mass Cytometry (CyTOF) and Realtime Deformation of Cells (RT-DC).

While these methods produce many thousands of output points, many of the analyses recently published with these novel technologies take point estimates on central tendency measures (such as mean or median) of sparse data sets to estimate the statistical significance of effects. These estimates are likely to be underpowered and to overstate effects, (Vashishth, Gelman et al.), fail to accurately estimate uncertainty or correlations between parameters (when they can even be estimated), and take no advantage of the distribution of thousands of data points available with each measurement.

To optimally address these research questions, we developed an up-to-date, Bayesian statistical approach that can be applied to both of these new cell-measurement technologies as well as others. This approach incorporates quantile summary statistics to more fully estimate the distribution of individual measurements; applies orthogonal contrast codings; handles more fully specified models even with relatively sparse data, and produces a probability mass that enables posterior evaluation of the uncertainty of parameter estimates. Particularly with novel technologies and pilot studies, evaluation of uncertainty in parameter estimates is key to robust and reproducible research, and we provide a method of delivering both estimate and uncertainty that is straightforward to use, code and interpret, that may be useful for a wide range of biological measurement tools.

CyTOF

Time-Of-Flight Mass Cytometry (CyTOF) measures spectral response of different cell types at the cellular level. Here the research question was whether there were differences in performance between three clinically used Gadolinium compounds: Gadovist, Magnevist and Dotarem. As these three compounds have different safety levels, investigating whether the signal levels were also different could aid clinical policy in choosing which compound to use.

Samples were analyzed for each of six cell types. In this pilot study a single subject was analyzed three times at three different concentration levels.

```
##    Subject              Date           Experiment   Contrast_Agent
##  Length:180         Length:180         1:60       control  :18
##  Class :character   Class :character   2:60       Dotarem  :54
##  Mode  :character   Mode  :character   3:60       Gadovist :54
##                                                   Magnevist:54
##
##
##  Concentration        Cell_Type  Marker_Type       CV
##  Min.   :0.00   B cells    :30   GdCl3:180   Min.   : 40.40
```

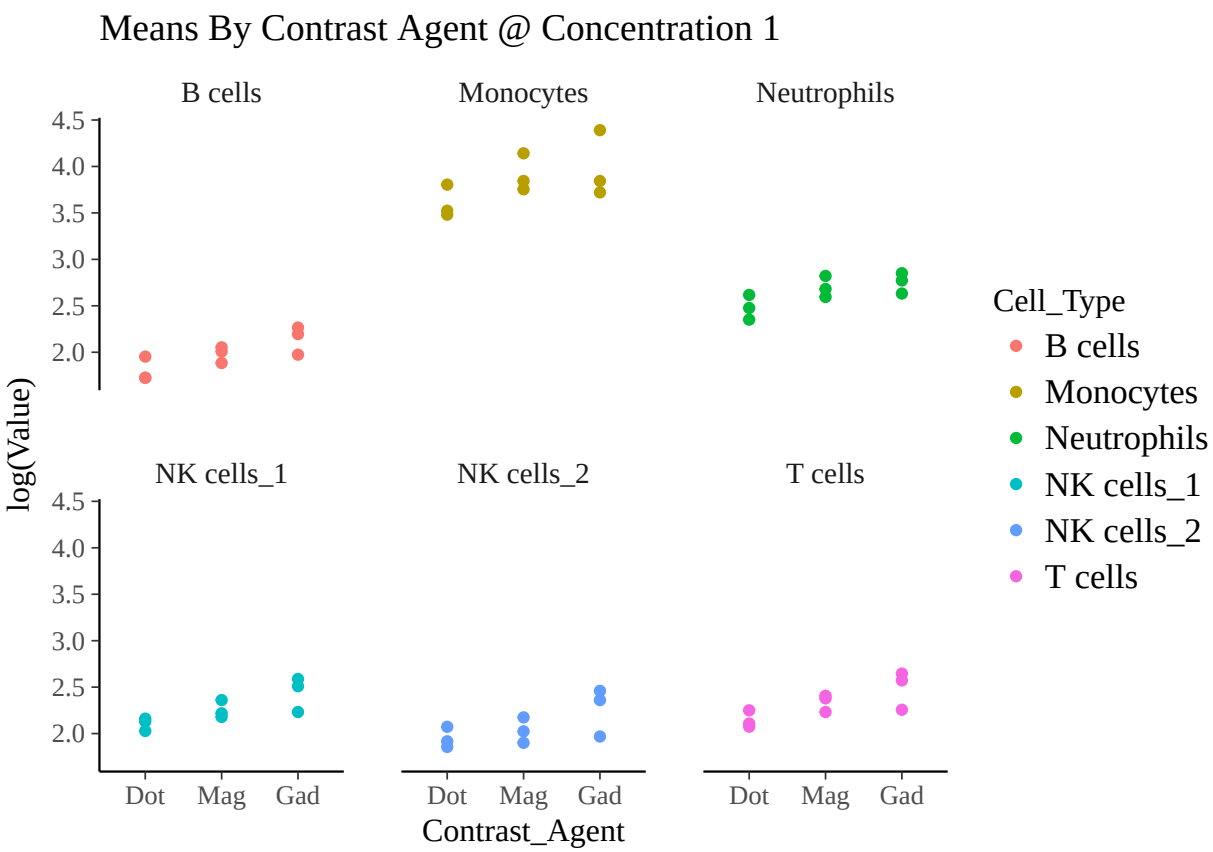
```
##  1st Qu.:0.10   Monocytes  :30                1st Qu.: 71.55
##  Median :0.30   Neutrophils:30                Median : 82.65
##  Mean   :0.42   NK cells_1 :30                Mean   : 84.48
##  3rd Qu.:1.00   NK cells_2 :30                3rd Qu.: 93.70
##  Max.   :1.00   T cells    :30                Max.   :273.00
##      Events              Mean            Median           Q05
##  Min.   :   334.8   Min.   : 0.760   Min.   : 0.580   Min.   : 0.2200
##  1st Qu.:  1577.4   1st Qu.: 2.027   1st Qu.: 1.500   1st Qu.: 0.4875
##  Median :  5222.0   Median : 3.760   Median : 2.955   Median : 0.7450
##  Mean   : 22134.3   Mean   : 7.679   Mean   : 6.343   Mean   : 2.0194
##  3rd Qu.: 43983.0   3rd Qu.: 9.033   3rd Qu.: 7.490   3rd Qu.: 1.7575
##  Max.   :102819.6   Max.   :80.700   Max.   :67.900   Max.   :30.4000
##       Q95          logConcentration    logMean          SD_conserv
##  Min.   :  1.570   Min.   :  -Inf   Min.   :-0.2744   Min.   :0.3289
##  1st Qu.:  5.018   1st Qu.:-2.303   1st Qu.: 0.7068   1st Qu.:0.8104
##  Median :  9.510   Median :-1.204   Median : 1.3244   Median :0.8864
##  Mean   : 17.647   Mean   :  -Inf   Mean   : 1.4708   Mean   :0.8817
##  3rd Qu.: 21.100   3rd Qu.: 0.000   3rd Qu.: 2.2008   3rd Qu.:0.9630
##  Max.   :181.000   Max.   : 0.000   Max.   : 4.3907   Max.   :1.1325
##        SE
##  Min.   :0.003168
##  1st Qu.:0.004258
##  Median :0.011354
##  Mean   :0.014392
##  3rd Qu.:0.019176
##  Max.   :0.056542
```

Our specific statistical question is whether there is an effect of contrast agent on the signal, independent of trial and cell type. It is clear that results will covary with cell type, but we expect the distribution around the "experiment" parameter to be iid.

Visualizing the means

Here we plot means at concentration 1 for each contrast agent within each cell type. There seems good evidence of an effect where Dot < Mag and Mag < Gad. We use this sliding contrast coding.

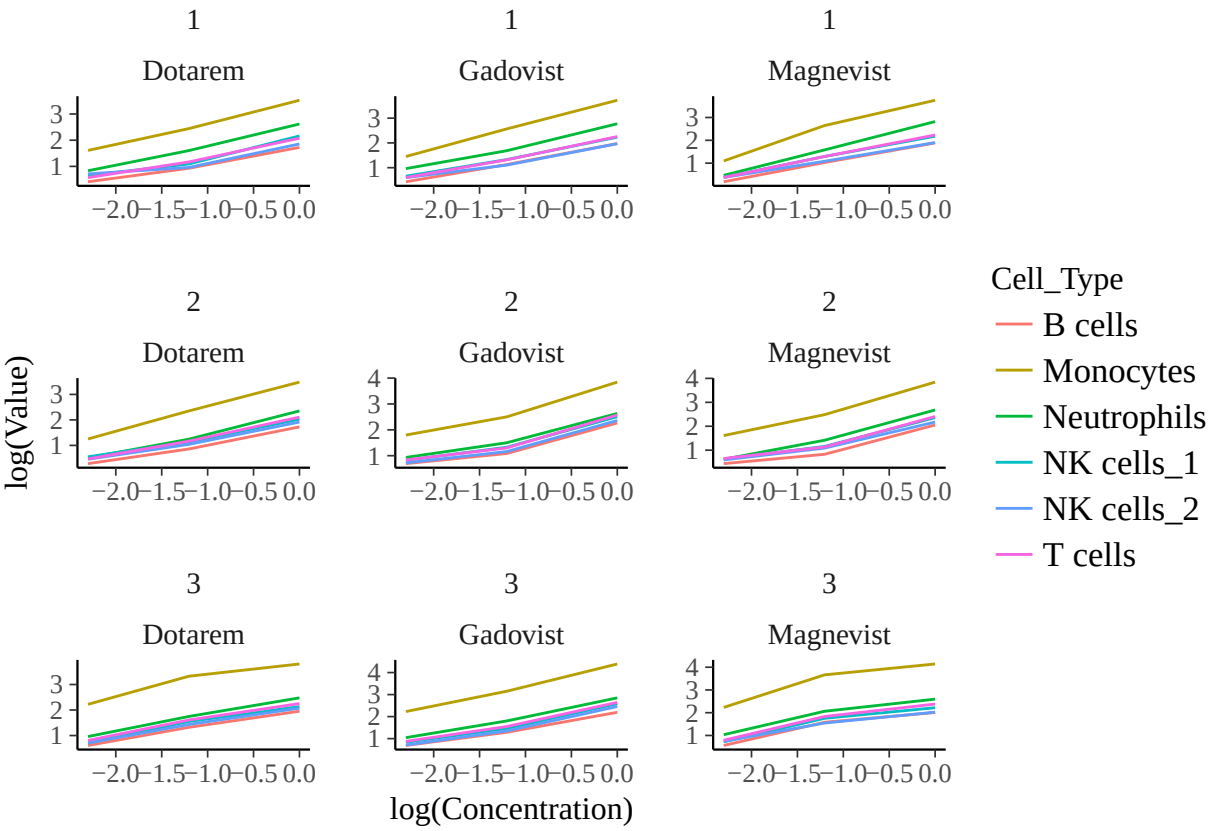
```
means <- subset(masscyto_data_tall, (Stat_Type == "Mean") & Contrast_Agent != "control")
means_conc_1 <- subset(means, (Concentration == 1))
means_conc_1$Contrast_Agent <- factor(means_conc_1$Contrast_Agent, c("Dotarem", "Magnevist", "Gadovist"))
plt <- ggplot(means_conc_1) +
    geom_point(aes(x=Contrast_Agent, y=log(Value), color=Cell_Type)) +
    facet_wrap(~ Cell_Type, ncol=3) +
    ggtitle("Means By Contrast Agent @ Concentration 1") +
    scale_x_discrete(labels = c("Dot", "Mag", "Gad"))
print(plt)
```



Log-log means plot

The log-lot plot looks reasonably linear:

```
ggplot(subset(means, Contrast_Agent != 'control')) +
    geom_line(aes(x=log(Concentration), y=log(Value), group=Cell_Type, color=Cell_Type)) +
    facet_wrap(~ Experiment + Contrast_Agent, scales='free')
```



In contrast to the previous normal model, which evaluated slopes, here we model the slope as uniform and the change as occurring in the *intercept* of the data. Consequently we will interpret change in *intercept* as an effect. The model is as follows:

SV: I don't understand what you mean about the intercept above. In the model below, it's still the slopes for GvsM and MvsD (and their interaction with concentration) that we are interpreting, isn't it? Oh, I just understood what you are saying here. It's actually the slopes that we are interpreting here; maybe what is confusing about brms is that all *coefficients* (called b in the priors) have a Cauchy(0,10) prior. So, we should rather say:

We are again interpreting the slopes: the effect of GvsM and MvsD. A positive value for the respective parameter means that G > M and M > D.

```
## hand-coded sliding contrasts:
masscyto_data_wide$GvsM<-ifelse(masscyto_data_wide$Contrast_Agent=="Gadovist",1,
           ifelse(masscyto_data_wide$Contrast_Agent=="Magnevist",-1,0))
masscyto_data_wide$MvsD<-ifelse(masscyto_data_wide$Contrast_Agent=="Magnevist",1,
           ifelse(masscyto_data_wide$Contrast_Agent=="Dotarem",-1,0))

priors<-c(set_prior("cauchy(0,10)", class = "b"),
          set_prior("normal(0,1)", class = "sd"),
          set_prior("lkj(2)", class = "cor"))

brm_df <- subset(masscyto_data_wide[c("logMean", "SE", "logConcentration", "GvsM", "MvsD", "Cell_Type",
brm_df$Contrast_Agent <- factor(brm_df$Contrast_Agent, labels = c("Gadovist", "Magnevist", "Dotarem"))

cytof_brm_log<-brm(formula = logMean | se(SE) ~
            logConcentration +
            GvsM+
```

```
            MvsD+
            logConcentration:GvsM +
            logConcentration:MvsD +
                (1 +
                logConcentration +
                GvsM+
                MvsD+
                logConcentration:GvsM +
                logConcentration:MvsD
                | Cell_Type),
            data = brm_df,
            family = gaussian(),
            prior = priors,
            warmup = 1000,
            iter = 2000,
            chains = 4,
            control = list(adapt_delta = 0.99,max_treedepth=15))
```

```
## Compiling the C++ model
```

```
## Start sampling
```

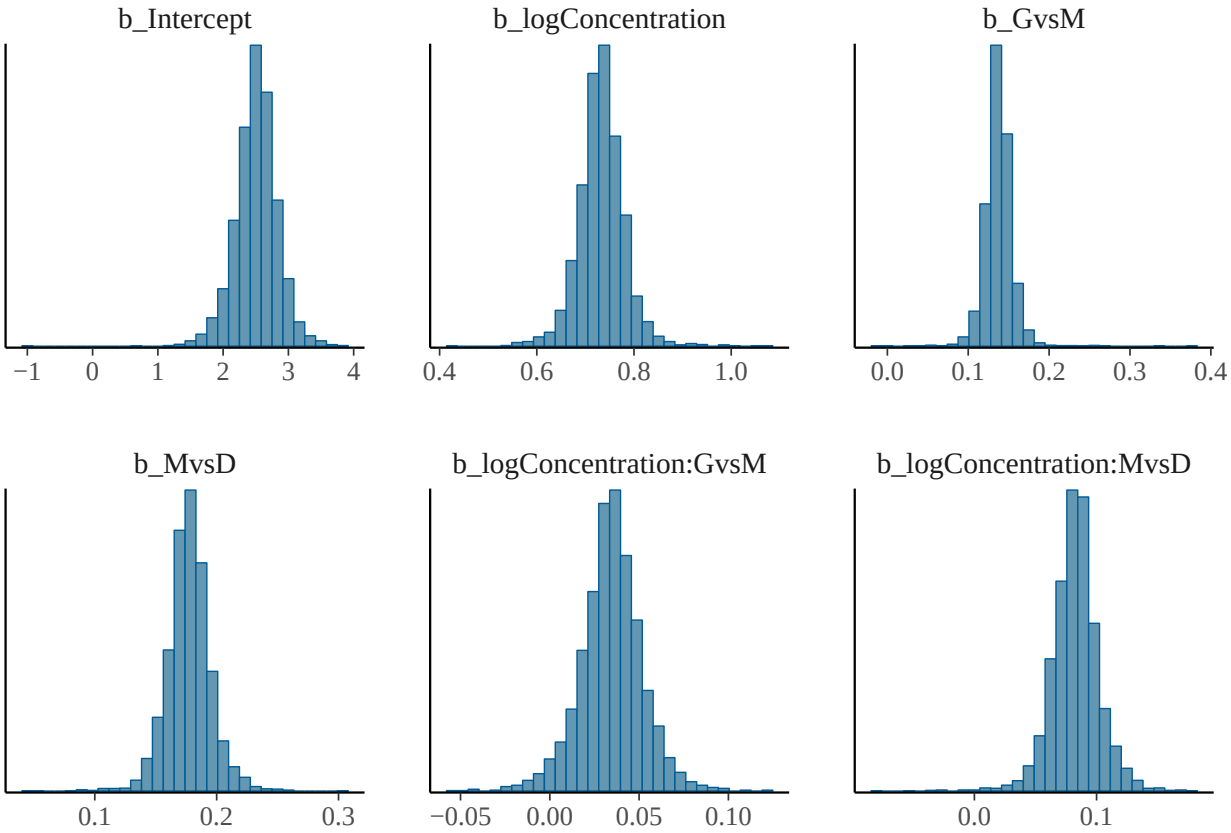
```
##
## SAMPLING FOR MODEL 'gaussian brms-model' NOW (CHAIN 1).
##
## Gradient evaluation took 0.000136 seconds
## 1000 transitions using 10 leapfrog steps per transition would take 1.36 seconds.
## Adjust your expectations accordingly!
##
##
## Iteration:    1 / 2000 [  0%]  (Warmup)
## Iteration:  200 / 2000 [ 10%]  (Warmup)
## Iteration:  400 / 2000 [ 20%]  (Warmup)
## Iteration:  600 / 2000 [ 30%]  (Warmup)
## Iteration:  800 / 2000 [ 40%]  (Warmup)
## Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Iteration: 2000 / 2000 [100%]  (Sampling)
##
##  Elapsed Time: 877.635 seconds (Warm-up)
##                899.955 seconds (Sampling)
##                1777.59 seconds (Total)
##
##
## SAMPLING FOR MODEL 'gaussian brms-model' NOW (CHAIN 2).
##
## Gradient evaluation took 9.1e-05 seconds
## 1000 transitions using 10 leapfrog steps per transition would take 0.91 seconds.
## Adjust your expectations accordingly!
##
##
```

```
## Iteration:    1 / 2000 [  0%]  (Warmup)
## Iteration:  200 / 2000 [ 10%]  (Warmup)
## Iteration:  400 / 2000 [ 20%]  (Warmup)
## Iteration:  600 / 2000 [ 30%]  (Warmup)
## Iteration:  800 / 2000 [ 40%]  (Warmup)
## Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Iteration: 2000 / 2000 [100%]  (Sampling)
##
##  Elapsed Time: 649.841 seconds (Warm-up)
##                801.129 seconds (Sampling)
##                1450.97 seconds (Total)
##
##
## SAMPLING FOR MODEL 'gaussian brms-model' NOW (CHAIN 3).
##
## Gradient evaluation took 8.3e-05 seconds
## 1000 transitions using 10 leapfrog steps per transition would take 0.83 seconds.
## Adjust your expectations accordingly!
##
##
## Iteration:    1 / 2000 [  0%]  (Warmup)
## Iteration:  200 / 2000 [ 10%]  (Warmup)
## Iteration:  400 / 2000 [ 20%]  (Warmup)
## Iteration:  600 / 2000 [ 30%]  (Warmup)
## Iteration:  800 / 2000 [ 40%]  (Warmup)
## Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Iteration: 2000 / 2000 [100%]  (Sampling)
##
##  Elapsed Time: 825.844 seconds (Warm-up)
##                975.931 seconds (Sampling)
##                1801.78 seconds (Total)
##
##
## SAMPLING FOR MODEL 'gaussian brms-model' NOW (CHAIN 4).
##
## Gradient evaluation took 0.000139 seconds
## 1000 transitions using 10 leapfrog steps per transition would take 1.39 seconds.
## Adjust your expectations accordingly!
##
##
## Iteration:    1 / 2000 [  0%]  (Warmup)
## Iteration:  200 / 2000 [ 10%]  (Warmup)
## Iteration:  400 / 2000 [ 20%]  (Warmup)
## Iteration:  600 / 2000 [ 30%]  (Warmup)
```

```
## Iteration:  800 / 2000 [ 40%]  (Warmup)
## Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Iteration: 2000 / 2000 [100%]  (Sampling)
##
##  Elapsed Time: 843.49 seconds (Warm-up)
##                622.2 seconds (Sampling)
##                1465.69 seconds (Total)
```

```
stanplot(cytof_brm_log, type="hist",pars=c("^b"))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



We then want to view the exponent of the posterior distributions, this time interested in the intercept:

```
## [1] "95% HDI, log(Gad) > log(Mag) :"
```

```
##    lower    upper
## 1.112265 1.185660
## attr(,"credMass")
## [1] 0.95
```

```
## [1] "95% HDI, log(Mag) > log(Dot) :"
```

```
##    lower    upper
## 1.149736 1.236173
## attr(,"credMass")
```

```
## [1] 0.95
```

SV: If our model is

$log(value) = \beta_1 + \beta_2 GvsM \ldots$

Shouldn't the transformation be

$exp(\beta_1 + \beta_2)$ and not just $exp(\beta_2)$?

```
mag_gad_log <-unlist(exp(posterior_samples(cytof_brm_log,'b_Intercept',exact_match=TRUE)+
                         posterior_samples(cytof_brm_log, 'b_GvsM',exact_match=TRUE)))
print('95% credible interval, log(Gad) > log(Mag) :')
```

```
## [1] "95% credible interval, log(Gad) > log(Mag) :"
```

```
print(hdi(mag_gad_log))
```

```
##     lower     upper
##  6.050203 24.138360
## attr(,"credMass")
## [1] 0.95
```

SV: Similarly for the other case. Also, you wrote Highest (posterior) Density Interval, but this is a credible interval. The two are the same if the distribution is symmetric, but if not then they differ. The credible interval is an equal-tailed interval; the HDI not necessarily.

In the log case the relationship is multiplicative, so if 95% of the probability mass is greater than 1, there is a 95% confident effect that the contrast agent increases the signal.

SV: Is this what you mean (see code immediately below)? How can probability mass be greater than 1? I didn't understand that statement. I would just say: we can compute the posterior probability that the parameter has a positive sign. If it is positive, it means that the contrast agent G increases the signal *more* than M (the sliding contrasts do pairwise comparisons).

```
mean(unlist(posterior_samples(cytof_brm_log, 'b_GvsM',exact_match=TRUE))>0)
```

```
## [1] 0.99975
```

Predictive modelling

While the goal of the analysis was to find the impact of the contrast agent independent of cell type and experiment, it leaves the result somewhat abstract. We can use the model to predict, for example, what impact a change of contrast agent will have on each cell types. For example we can ask the question, if we use Gadovist instead of Magnevist, what is the expected increase in signal for each unit of concentration, as estimated at concentration 1:

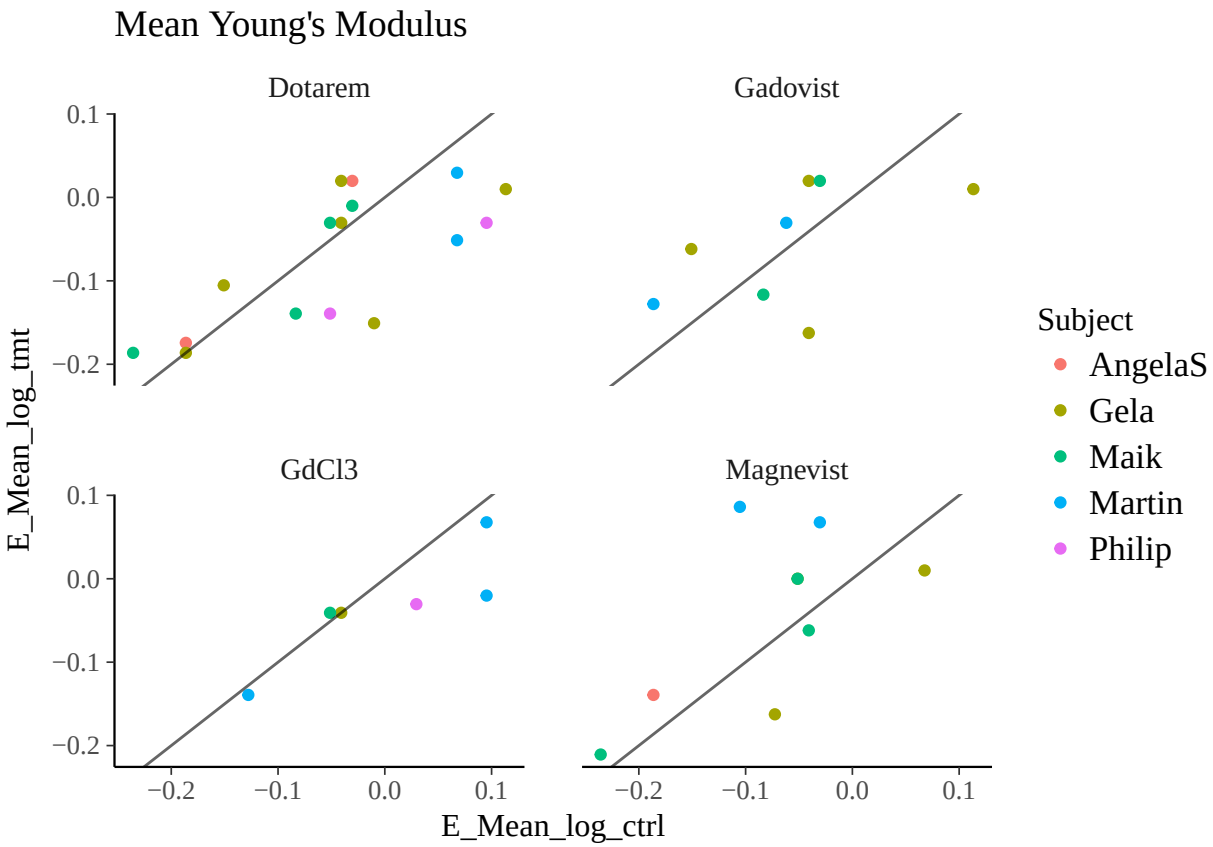
RT-DC

RT-DC is a method of measuring cellular elastic deformations developed by TU-Dresden. We use this technology to evaluate whether cells containing common contrast agents show differences in the Young's modulus, which would indicate that their physiological properties are being altered by the contrast agent.

Distances between mean and 05 quantile are much smaller than distances between mean and 95 quantile, but in the log transform they are quite similar. So again we use a log model (in this case, there is no concentration).

```
##   Subject   Date Cell_Type Contrast_Agent  Protocol Events E_Mean E_SD
## 1    Gela 030117 Monocytes        Dotarem   Control    499   0.96 0.16
## 2    Gela 030117 Monocytes        Dotarem Treatment    603   0.97 0.18
## 3    Maik 030217 Monocytes        Dotarem   Control    426   0.95 0.16
## 4    Maik 030217 Monocytes        Dotarem Treatment    137   0.97 0.28
## 5 AngelaS 030317 Monocytes        Dotarem   Control    727   0.97 0.17
## 6 AngelaS 030317 Monocytes        Dotarem Treatment    826   1.02 0.17
##   E_q05 E_q95 E_SD_hi_log E_SD_low_log E_SE_conserv_log  E_Mean_log
## 1  0.76  1.23   0.2478362    0.2336149      0.011094670 -0.04082199
## 2  0.76  1.23   0.2374734    0.2439776      0.009670664 -0.03045921
## 3  0.75  1.19   0.2252466    0.2363888      0.010913234 -0.05129329
## 4  0.73  1.39   0.3597630    0.2842515      0.030736624 -0.03045921
## 5  0.74  1.25   0.2536028    0.2706459      0.009405606 -0.03045921
## 6  0.77  1.33   0.2653763    0.2811674      0.009233623  0.01980263
```

The RT-DC method is currently experimental, and control tests for an individual subject produce a wide range of values. Consequently each measurement with a contrast agent was paired with its own control from the same session. The Figure below shows pairwise control-treatment plots colored by subject:




```
##    Subject   Date   Cell_Type Contrast_Agent E_diff_Mean   E_diff_SE
## 1  AngelaS 030317   Monocytes        Dotarem -0.05026183 0.013172732
## 2  AngelaS 030317 Neutrophils        Dotarem -0.01197619 0.009318538
## 3  AngelaS 030317 Neutrophils      Magnevist -0.04706751 0.010218979
## 4     Gela 030117   Monocytes        Dotarem -0.01036279 0.014622681
## 5     Gela 030117   Monocytes          GdCl3  0.00000000 0.015523972
## 6     Gela 030117 Neutrophils        Dotarem  0.00000000 0.015573092
## 7     Gela 061317   Monocytes        Dotarem -0.06062462 0.026909706
## 8     Gela 061317   Monocytes       Gadovist -0.06062462 0.021129609
```

```
## 9     Gela 061317   Monocytes      Magnevist -0.05129329 0.023606282
## 10    Gela 061317 Neutrophils        Dotarem -0.04546237 0.016885436
## 11    Gela 061317 Neutrophils       Gadovist -0.08894749 0.014653119
## 12    Gela 090417   Monocytes        Dotarem  0.10337835 0.007379648
## 13    Gela 090417   Monocytes       Gadovist  0.10337835 0.005996296
## 14    Gela 090417   Monocytes      Magnevist  0.05770832 0.007137553
## 15    Gela 090417 Neutrophils        Dotarem  0.14077255 0.011720506
## 16    Gela 090417 Neutrophils       Gadovist  0.12169693 0.015468479
## 17    Gela 090417 Neutrophils      Magnevist  0.08994824 0.014228503
## 18    Maik 030217   Monocytes        Dotarem -0.02083409 0.025298601
## 19    Maik 030217   Monocytes          GdCl3 -0.01047130 0.019872728
## 20    Maik 030217   Monocytes      Magnevist -0.05129329 0.019020734
```

```
##   Subject   Date Cell_Type Contrast_Agent E_diff_Mean  E_diff_SE Gad_Mag
## 1 AngelaS 030317 Monocytes        Dotarem -0.05026183 0.01317273       0
## 4    Gela 030117 Monocytes        Dotarem -0.01036279 0.01462268       0
## 5    Gela 030117 Monocytes          GdCl3  0.00000000 0.01552397       0
## 7    Gela 061317 Monocytes        Dotarem -0.06062462 0.02690971       0
## 8    Gela 061317 Monocytes       Gadovist -0.06062462 0.02112961      -1
## 9    Gela 061317 Monocytes      Magnevist -0.05129329 0.02360628       1
##   Mag_Dot  GdCl3_All
## 1       1  0.3333333
## 4       1  0.3333333
## 5       0 -1.0000000
## 7       1  0.3333333
## 8       0  0.3333333
## 9      -1  0.3333333
```

Monocytes

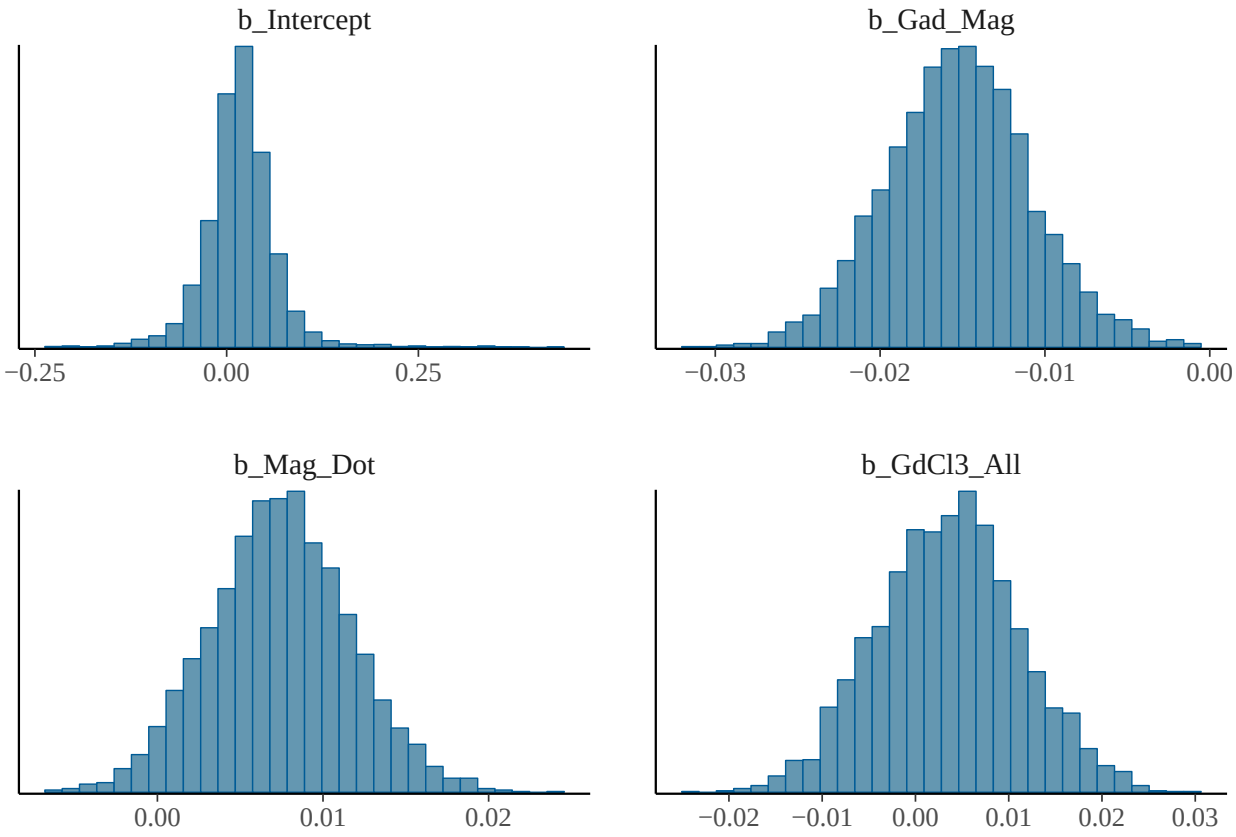
```
priors_cauchy_diff2 <- c(set_prior("cauchy(0, 10)", class = "Intercept"),
            set_prior("cauchy(0, 10)", class = "b"),
            set_prior("cauchy(0, 10)", class = "sd")
            )


mdiffMonoContrAgt_E_ME <- brm(formula = E_diff_Mean | se(E_diff_SE) ~ 1+Gad_Mag+Mag_Dot+GdCl3_All+
               (1 | Subject),
             data = diffMono, family = gaussian(), prior = priors_cauchy_diff2,
             iter = 2000, chains = 4, control = list(adapt_delta = 0.999))
```

```
## Compiling the C++ model
```

```
## Start sampling
```

```
stanplot(mdiffMonoContrAgt_E_ME, type="hist",pars=c("^b"))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
mMonoContrAgtdiffEpostME<-posterior_samples(mdiffMonoContrAgt_E_ME, "^b")
```

For Monocytes, there is weak evidence that GdCl3 causes a decrease in Young's Modulus against the other three contrast agents. The comparison of the individual contrast agents is not determinate at this time:

```
## [1] "Prob Gad > Mag"
## [1] 0
## [1] "Prob Gad < Mag"
## [1] 1
## [1] "Prob Mag > Dot"
## [1] 0.961
## [1] "Prob GdCl3 > All"
## [1] 0.675
```

Neutrophils

```
priors_cauchy_diff2 <- c(set_prior("cauchy(0, 10)", class = "Intercept"),
            set_prior("cauchy(0, 10)", class = "b"),
            set_prior("cauchy(0, 10)", class = "sd")
            )


mdiffNeutroContrAgt_E_ME <- brm(formula = E_diff_Mean | se(E_diff_SE) ~ 1+Gad_Mag+Mag_Dot+GdCl3_All+
```

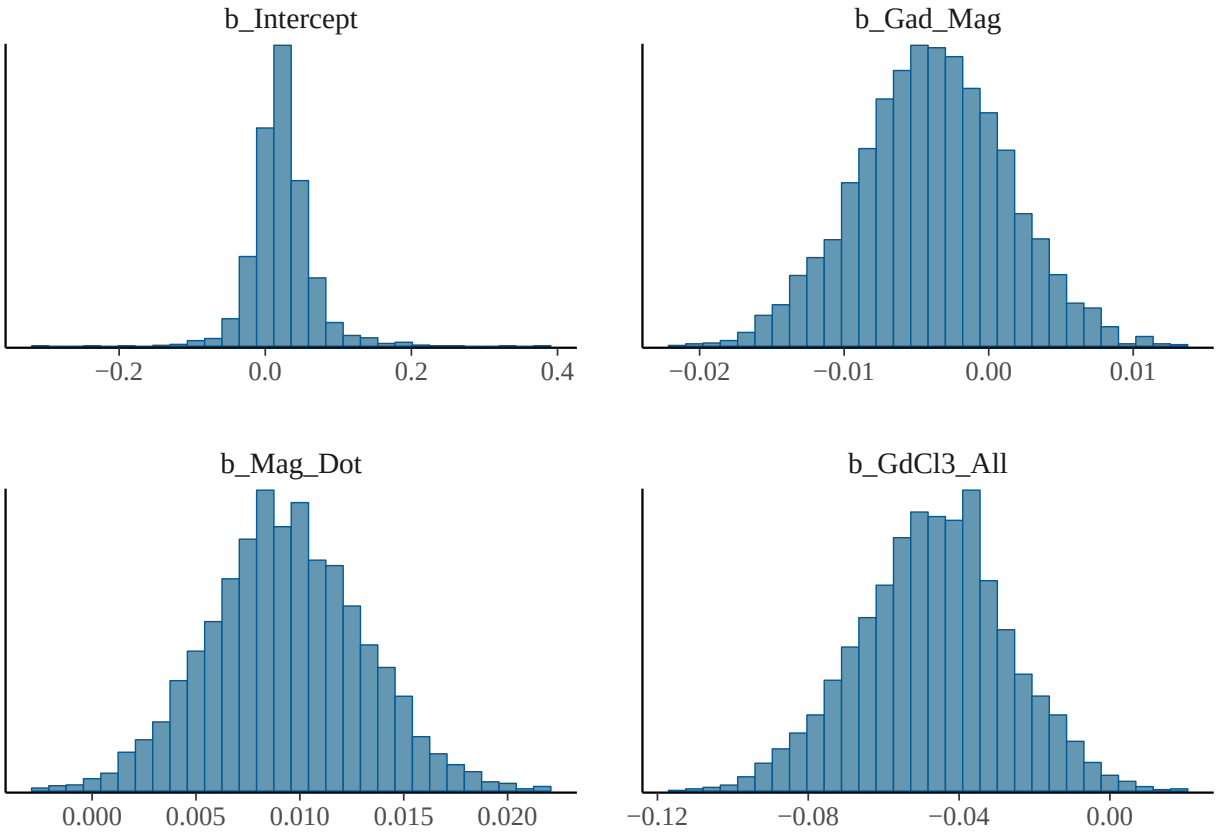
```
              (1 | Subject),
            data = diffNeutro, family = gaussian(), prior = priors_cauchy_diff2,
            iter = 2000, chains = 4, control = list(adapt_delta = 0.999))
```

```
## Compiling the C++ model
```

```
## Start sampling
```

```
stanplot(mdiffNeutroContrAgt_E_ME, type="hist",pars=c("^b"))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
mNeutroContrAgtdiffEpostME<-posterior_samples(mdiffNeutroContrAgt_E_ME, "^b")
```

For Neutrophils, there isn't much evidence for differences by Contrast Agent, though there is again maybe some weak evidence (80% likelihood) that Magnevist has higher Young's modulus than Dotarem:

```
## [1] "Prob Gad < Mag"
```

```
## [1] 0.7715
```

```
## [1] "Prob Mag > Dot"
```

```
## [1] 0.99375
```

```
## [1] "Prob GdCl3 < All"
```

```
## [1] 0.992
```