Initial Flu Data Analysis

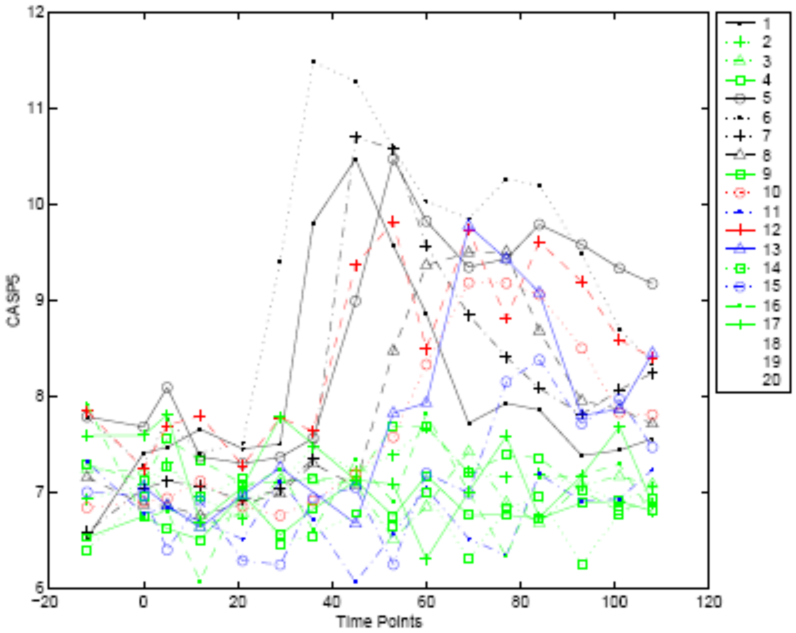
Ashish Tiwari, Mark-Oliver Stehr,
Carolyn Talcott, Patrick Lincoln

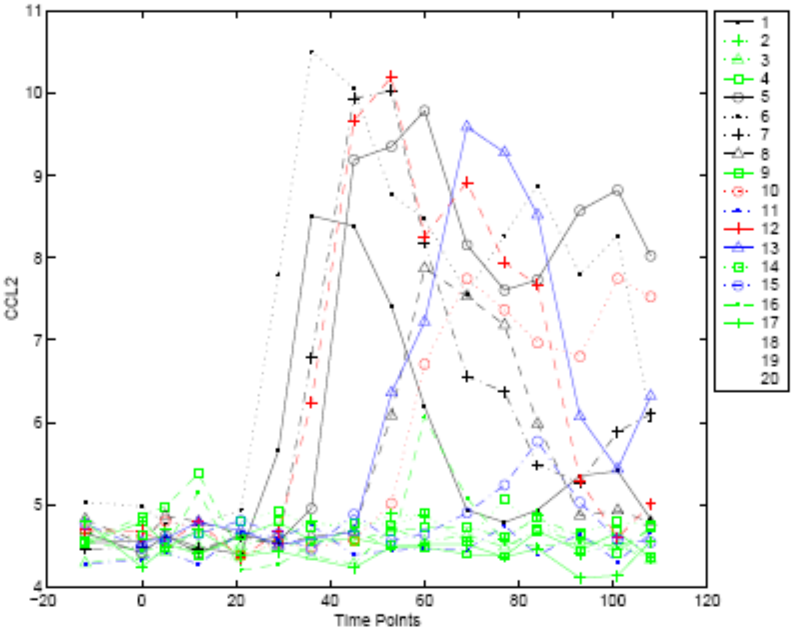
- Performed initial analysis of data
- Investigated baseline noise levels in gene expression data
- Developed pathway-based classifiers
- Performed now-usual leave-one-out cross validation
- Estimated statistical significance of results across all three diseases (rhino, rsv, flu)

Gene Expression

- Using set of genes identified as being associated with immune response to disease in early and middle phases:
  - CXCL10 OAS2 CCL2 CCL8 IL1B IL1RN CASP7 CCR1 CCR10 TAP1 TAP2 CASP5 TNFAIP6 TNFSF10

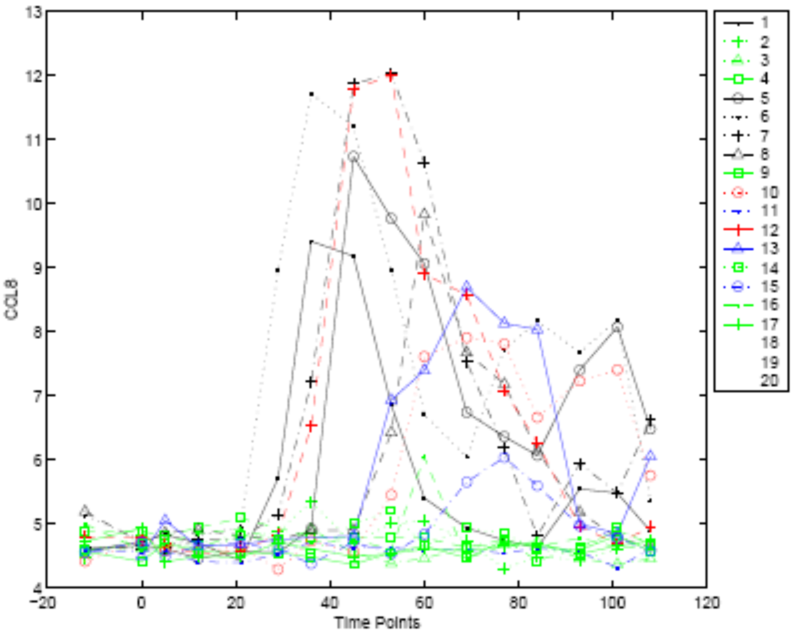
CASP5 and CCL2

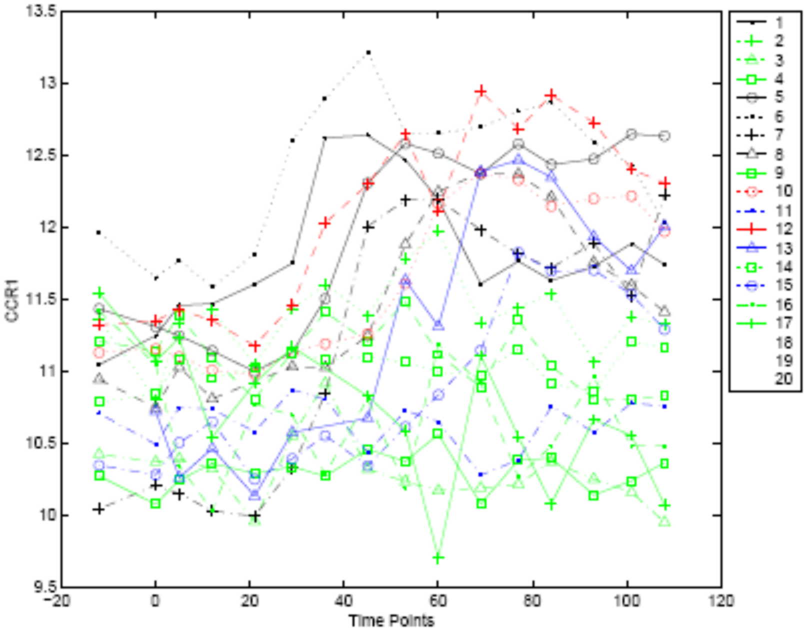




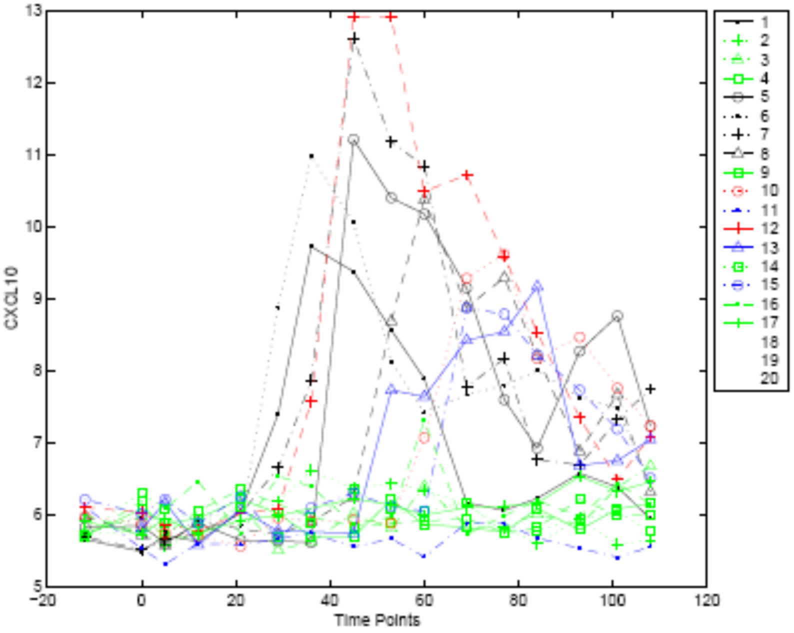
- Green: Asymptomatic
- Blue : Mildly symptomatic
- Red : Medium symptomatic
- Black : Strongly symptomatic

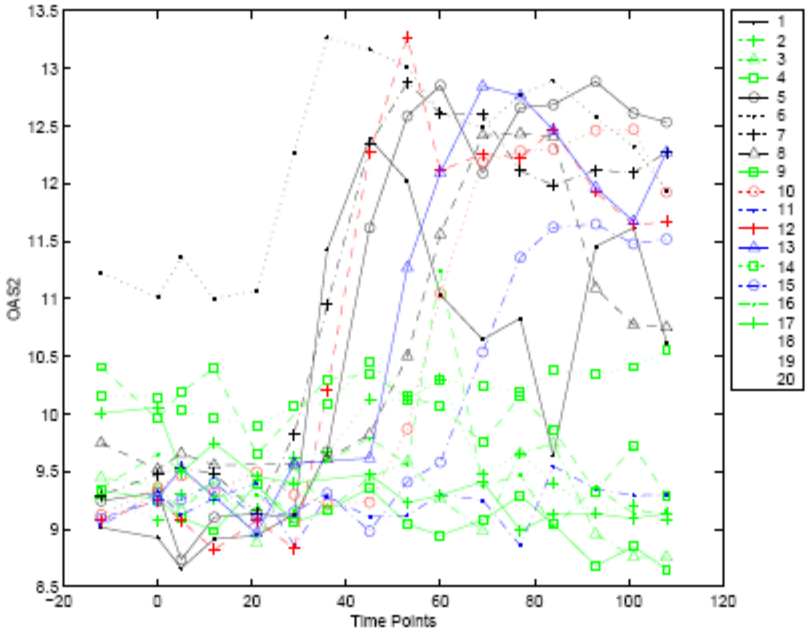
CCL8 and CCR1



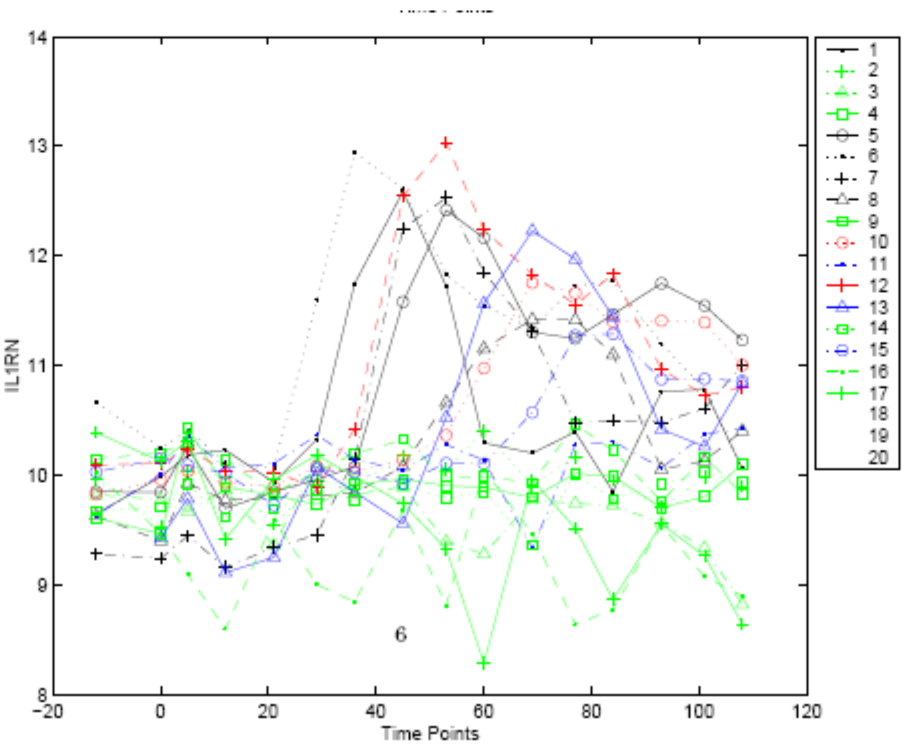


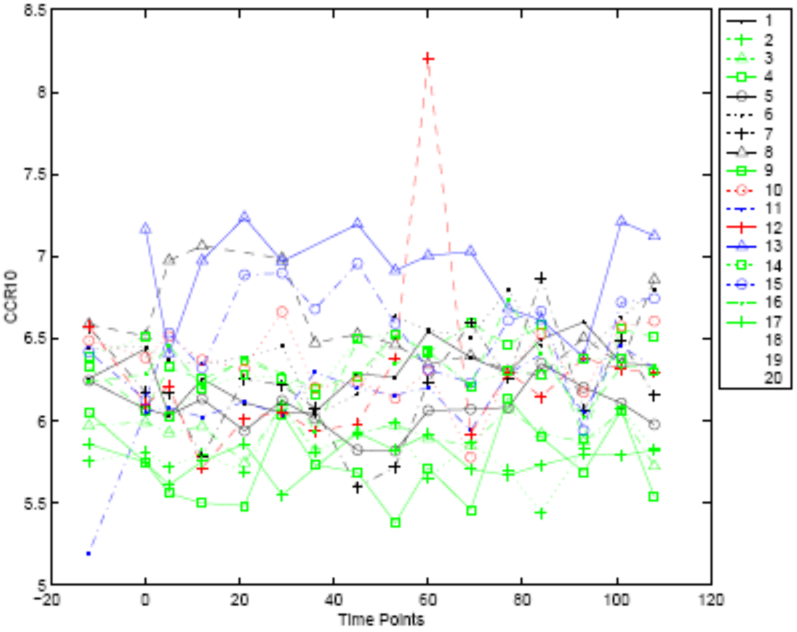
CXCL10 and OAS2





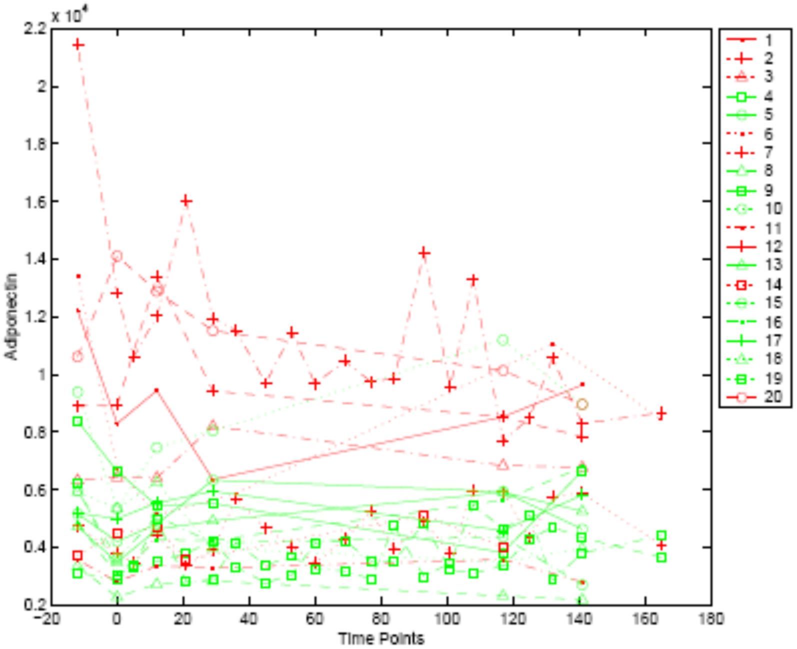
IL1RN and CCR10

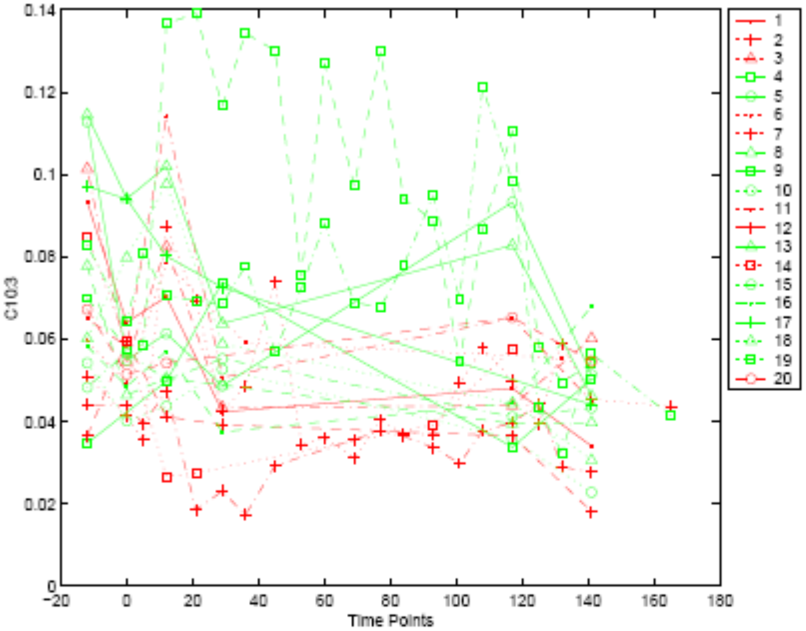




Metabolites

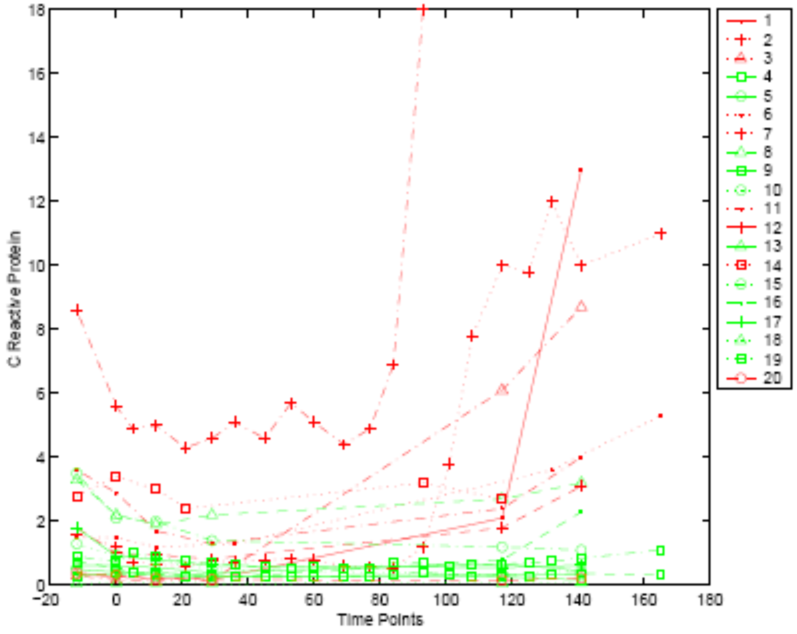
- Green: Asymptomatic
  - subjects: 4,5,8,9,10,13,15,16,17,18,19
- Red : Symptomatic
  - subjects: 1,2,3,6,7,11,12,14,20

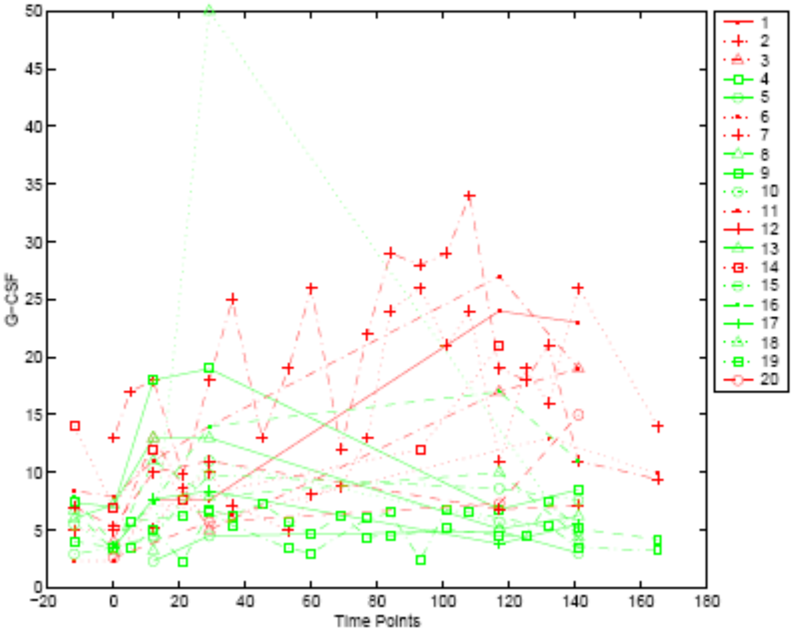




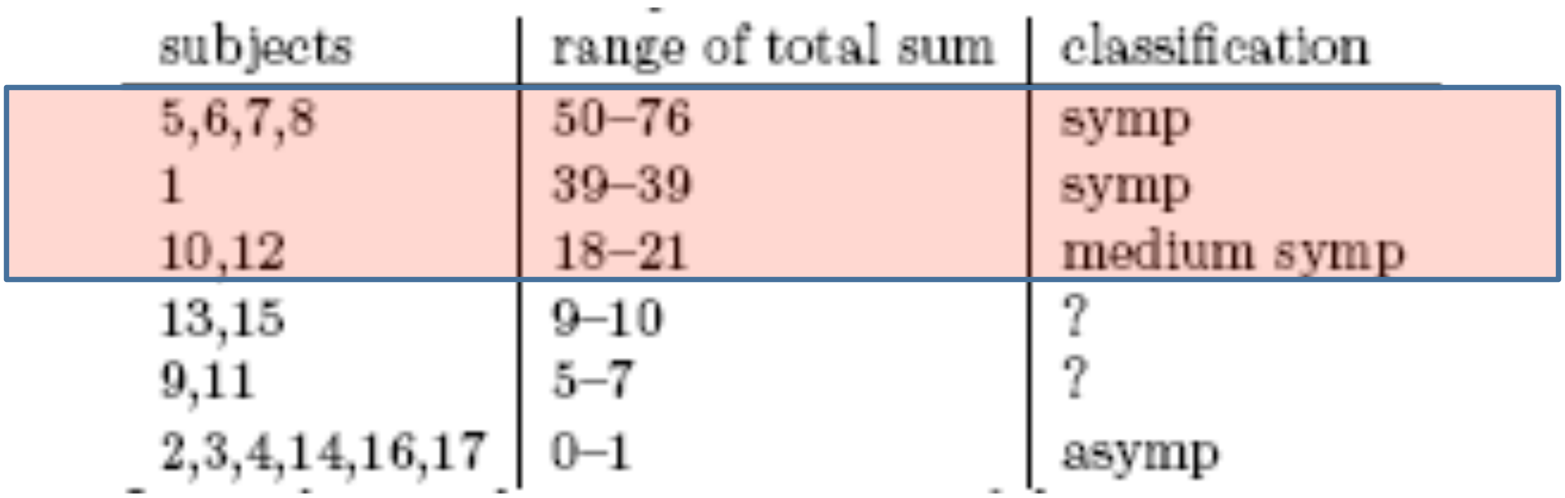
Proteomic Indicators

- Green: Asymptomatic
  - subjects: 4,5,8,9,10,13,15,16,17,18,19
- Red : Symptomatic
  - subjects: 1,2,3,6,7,11,12,14,20



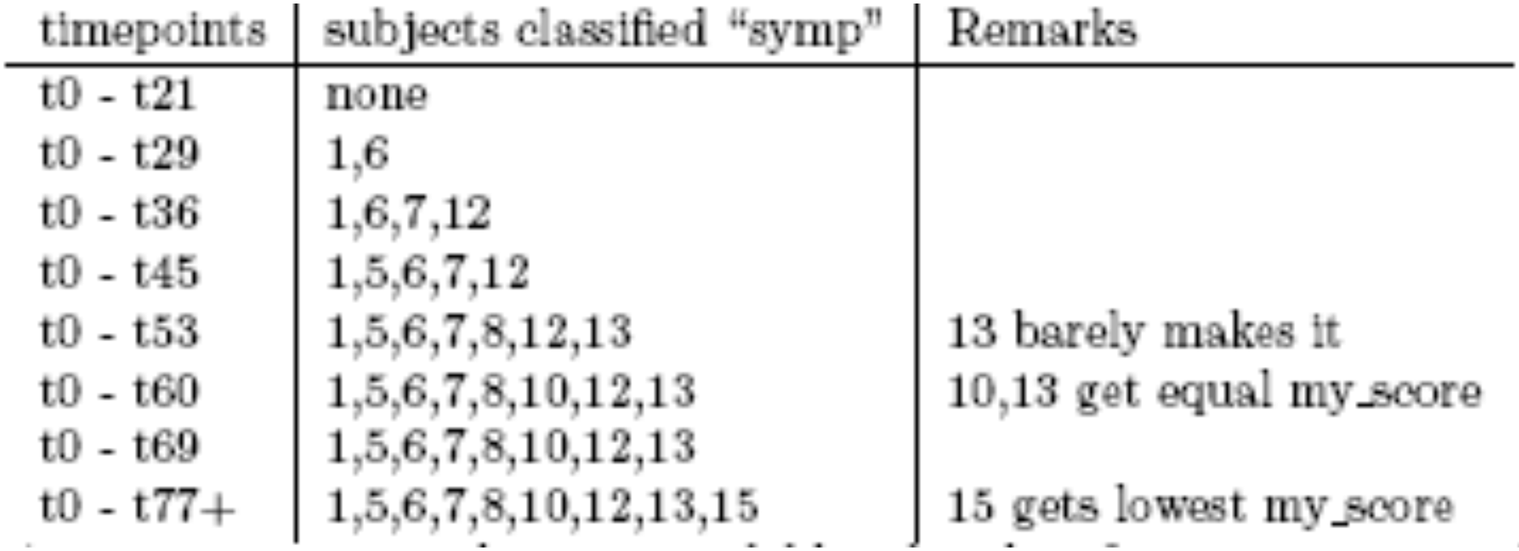


- Classification of subjects in flu study



subjects	range of total sum	classification
5,6,7,8	50–76	symp
1	39–39	symp
10,12	18–21	medium symp
13,15	9–10	?
9,11	5–7	?
2,3,4,14,16,17	0–1	asymp

Initial Classification Results: Strong signal in mid genes Subjects 13 and 15 appear sickly in late time points



timepoints	subjects classified "symp"	Remarks
t0 - t21	none	
t0 - t29	1,6	
t0 - t36	1,6,7,12	
t0 - t45	1,5,6,7,12	
t0 - t53	1,5,6,7,8,12,13	13 barely makes it
t0 - t60	1,5,6,7,8,10,12,13	10,13 get equal my_score
t0 - t69	1,5,6,7,8,10,12,13	
t0 - t77+	1,5,6,7,8,10,12,13,15	15 gets lowest my_score

Initial Conclusions

- Strong signal in middle and late periods
- As usual, gene expression is best, proteomics provide some signal, and metabolites provide relatively weak information.
- Subject classification of mildly symptomatic cases should be discussed further