sbv IMPROVER
SYSTEMS BIOLOGY VERIFICATION

www.sbvimprover.com

Diagnostic Signature Challenge: Perennial Benchmarking

Raquel Norel, PhD

Functional Genomics & Systems Biology, IBM Computational Biology Center

Diagnostic Signature Challenge: gene expression to phenotype prediction

Aim to assess and verify computational approaches that classify clinical samples based on transcriptomics data.

Participants were asked to establish predictive signatures on unlabeled gene expression data sets in 4 disease areas

- Chronic Obstructive Pulmonary Disease
- Multiple Sclerosis
- Lung Cancer
- Psoriasis

Scope & Rationale

Scope:

- Online automatic scoring of new predictions and display of ranking compared to final challenge results

Rationale:

- Benefit the scientific community by:
  - allowing scientists to benchmark their methods
  - motivating scientists to improve the methods developed in the challenge open phase
- Motivate participants to participate in other sbv IMPROVER challenges
- Stay in contact with the sbv IMPROVER community
- Establish standards for Diagnostic Signature models

We promised to do it…

EDITOR'S CORNER

Systems Biomedicine 1:4, e26325; October/November/December 2013; © 2013 Landes Bioscience

sbv IMPROVER Diagnostic Signature Challenge

Design and results

Kahn Rhrissorrakrai,[1,*] J Jeremy Rice,[1] Stephanie Boue,[2] Marja Talikka,[2] Erhan Bilal,[1] Florian Martin,[2] Pablo Meyer,[1] Raquel Norel,[1] Yang Xiang,[2] Gustavo Stolovitzky,[1] Julia Hoeng,[2] Manuel C Peitsch[2]

[1]IBM Computational Biology Center; Yorktown Heights, NY USA; [2]Philip Morris Products SA; Research and Development; Neuchâtel, Switzerland

deliver a truly autonomous classification procedure. Additionally, a web-based platform is being developed for perennial benchmarking, where future teams can test their methods and perform a self-assessment of their predictions. Though this system will no longer qualify as strictly blinded testing, and requires careful interpretation of scores, it will offer intermediate results and is a stable method for participants to benchmark themselves against current front-runners.

Perennial: what do we mean with that?

Full Definition of PERENNIAL

1. present at all seasons of the year
2. persisting for several years usually with new herbaceous growth from a perennating part <*perennial asters*>
3.
   a. persistent, enduring <*perennial favorites*>
   b. continuing without interruption : constant, perpetual <*the perennial quest for certainty*> <*a perennial student*>
   c. regularly repeated or renewed : recurrent <*death is a perennial literary theme*>

AN ENCYCLOPÆDIA BRITANNICA COMPANY

Merriam-Webster

m-w.com

http://www.merriam-webster.com/dictionary/perennial; accessed October 7th, 2013

Perennial: the self-assessment trap

- Different persons will use the evaluation tool for their assessment as a learning tool
- One user testing over and over again will end up overfitting
- For more about issues with self-evaluations, please see:

Molecular Systems Biology 7; Article number 537; doi:10.1038/msb.2011.70
Citation: *Molecular Systems Biology* 7: 537

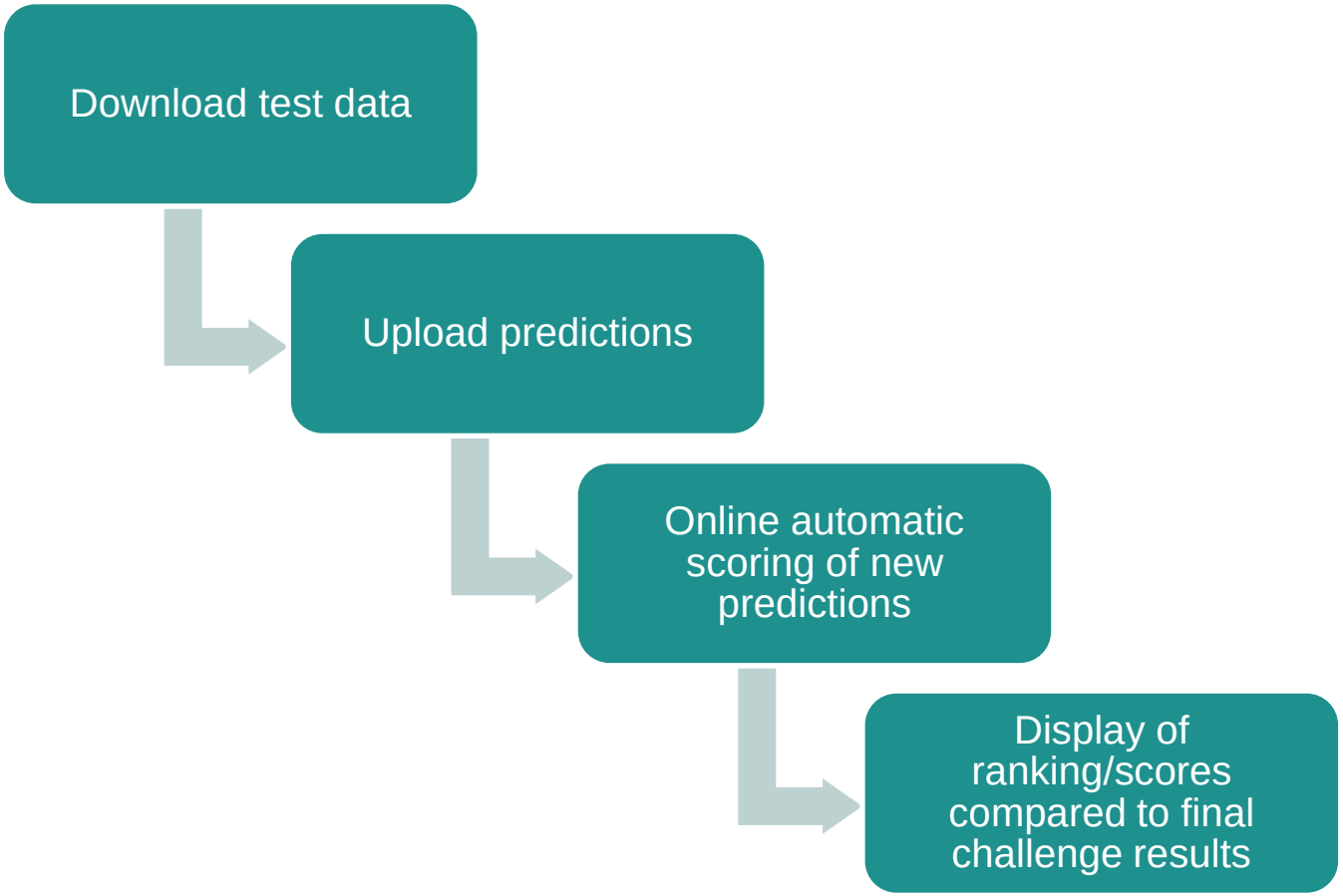
www.molecularsystemsbiology.com

CORRESPONDENCE

The self-assessment trap: can we all be better than average?

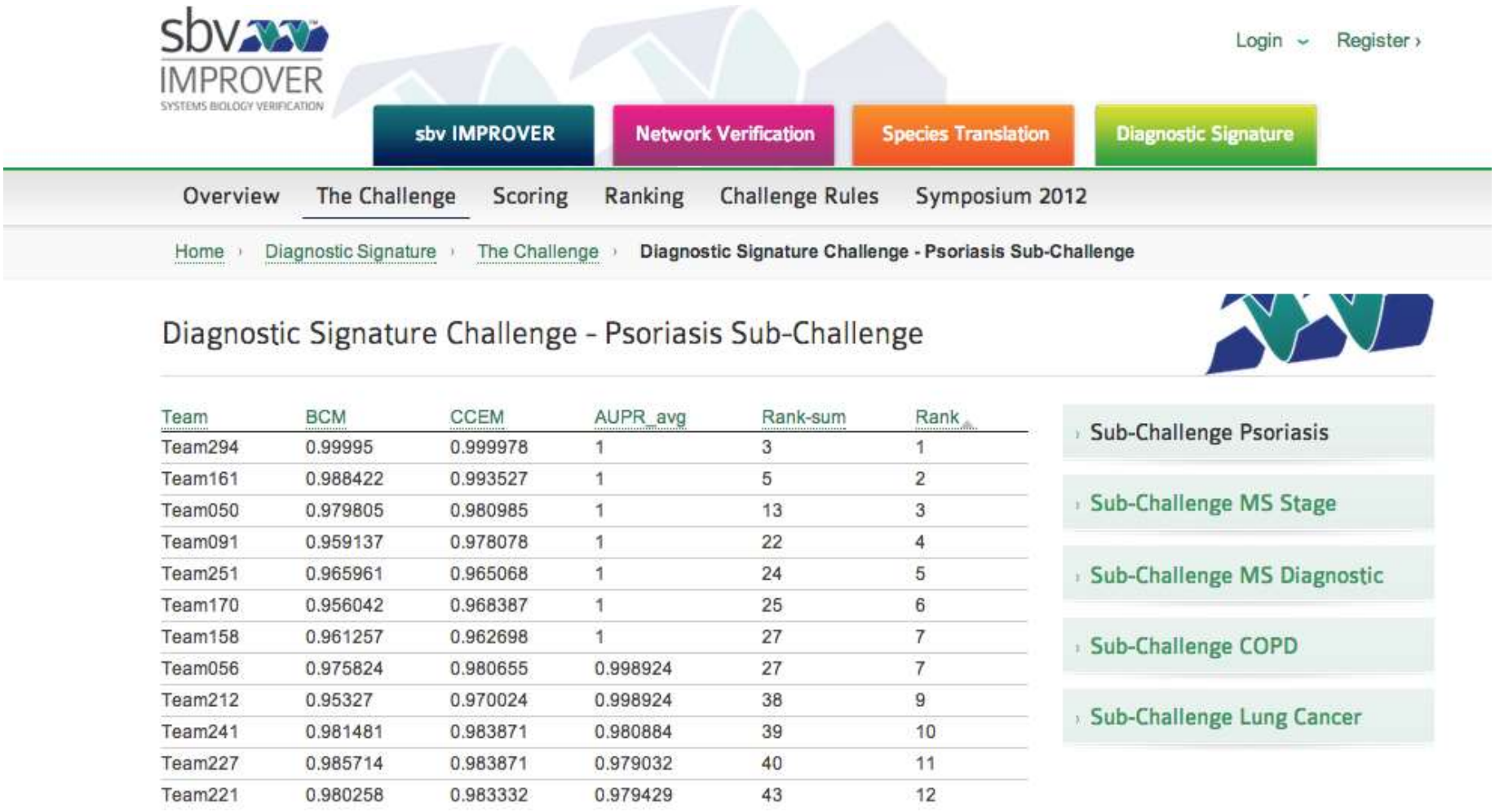
Molecular Systems Biology 7: 537; published online 11 October 2011; doi:10.1038/msb.2011.70

How can users expect to interact with the site?



We expect to include new data sets for testing your scripts, so stay tuned!

Sample screenshot



Team	BCM	CCEM	AUPR_avg	Rank-sum	Rank
Team294	0.99995	0.999978	1	3	1
Team161	0.988422	0.993527	1	5	2
Team050	0.979805	0.980985	1	13	3
Team091	0.959137	0.978078	1	22	4
Team251	0.965961	0.965068	1	24	5
Team170	0.956042	0.968387	1	25	6
Team158	0.961257	0.962698	1	27	7
Team056	0.975824	0.980655	0.998924	27	7
Team212	0.95327	0.970024	0.998924	38	9
Team241	0.981481	0.983871	0.980884	39	10
Team227	0.985714	0.983871	0.979032	40	11
Team221	0.980258	0.983332	0.979429	43	12

Thank you for your attention

The sbv IMPROVER project and www.sbvimprover.com are part of a collaboration designed to enable scientists to learn about and contribute to the development of a new crowd sourcing method for verification of scientific data and results. The project team includes scientists from Philip Morris International's (PMI) Research and Development department and IBM's Thomas J. Watson Research Center. The project is funded by PMI.