Text Mining Special Interest Group
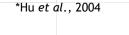
Stuart Murray, Wyeth Research

Novartis Institute for Biomedical Research, Cambridge, MA 6-8th October 2004

The PRISM Forum

At Wyeth Research, Text Mining is considered a solution for...?

- **Information Overload**

- To follow breast cancer a researcher would have to <u>SCAN</u> 130 different journals and <u>READ</u> 27 papers per day... ...in 1999*

- 1999 Breast cancer papers in Medline: 4589
- 2003 Breast Cancer papers in Medline: 5635

*Hu *et al.*, 2004

The **PRISM** Forum

What are the Critical Business Needs that Text Mining Could Address?

- Time management of scientists (Do more… …with less)

- Cross-therapeutic area exploitation

- Integration of information from core units
  – e.g. genomics, proteomics, informatics, literature

The **PRISM** Forum

What Questions do Users Want to Ask ?

- What are the normal biological roles of a target ?

- What are the normal interactions of a target ?

- What Key disease is a target involved in ?
  - What Key differences in roles or interactions of a target are known ?

The **PRISM** Forum

Text Mining Applications/Components

Function	Current	Future
Manual TM Methods	Keyword searching	-
3rd Party curation services	Ingenuity's IPA	-
Categorization & Clustering	Convera's Retrievalware	-
NLP or OBIIE	-	-
Full text access	Ovid	Ovid
Recall and precision metrics	-	-
Visualisation	Spotfire	-

Current Text Mining Applications in Wyeth Research

- OVID

- Refviz

- Integrity Pathway Analysis

- Web Resources: iHOP, MedGene, NCBI

The **PRISM** Forum

What Future Applications are Planned?

- To be determined...

The **PRISM** Forum

Conclusions

- Relational Databases – tools to provide a network as a natural way of accessing biological information...

- ...Add <u>VALUE</u> to target information and <u>VALUE</u> to information access

The **PRISM** Forum