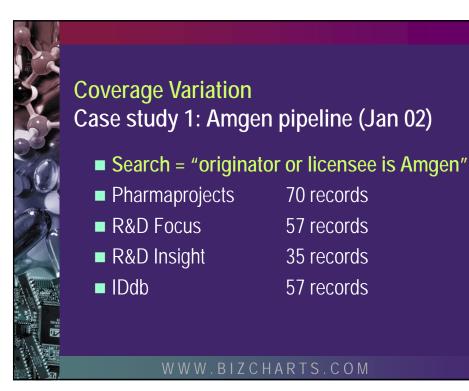


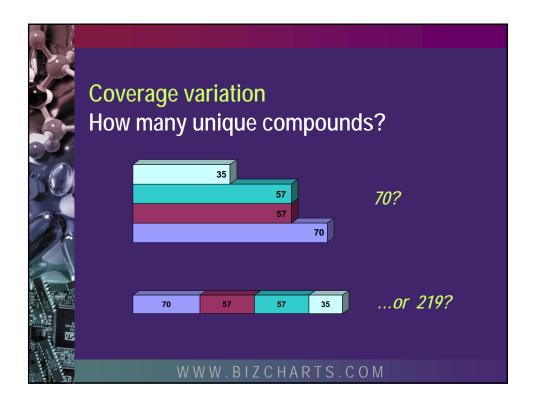


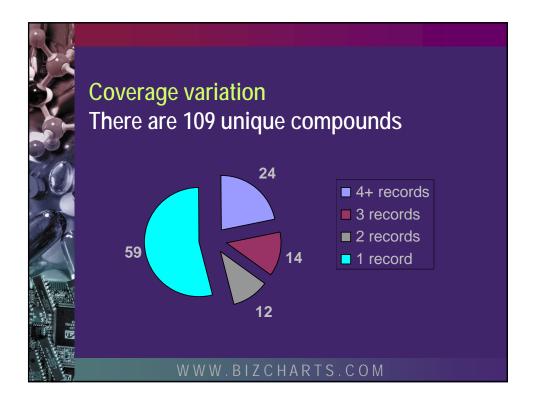


#### Drug Pipeline Databases Compare the differences in...

- Coverage the number of records retrieved from a particular search.
- Content -- the information provided for a compound by each database.









#### **Coverage Variation**

Case study 2: Epilepsy drugs (Aug 02)

- Search strategy: "Therapeutic activity code =N3A" (PP, RDF, RDI)
- "Indication = epilepsy" (IDdb)

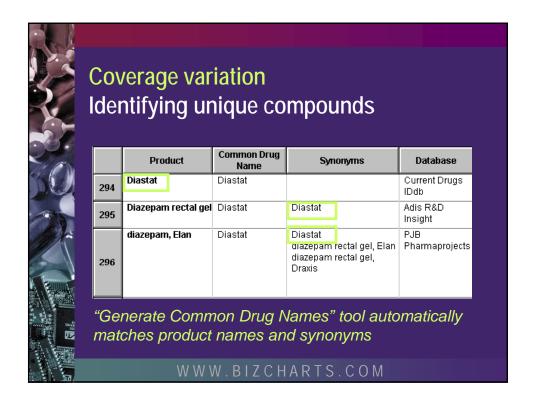
WWW BIZCHARTS COM

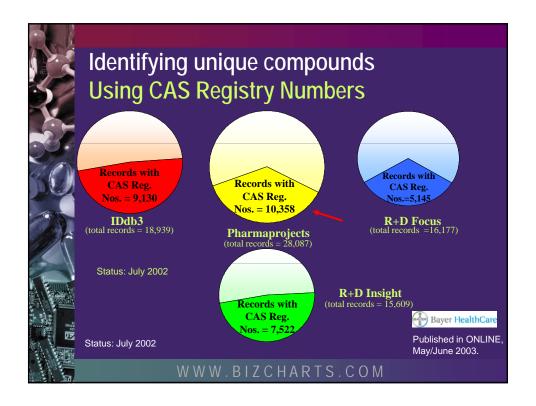


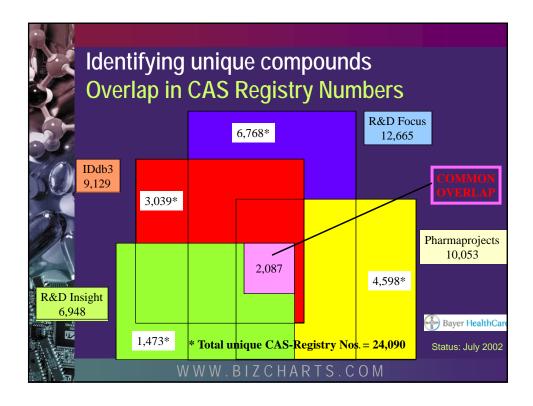
#### **Coverage Variation**

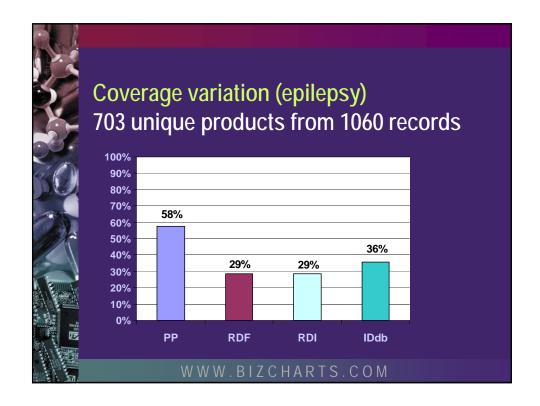
Case study 2: Epilepsy compounds

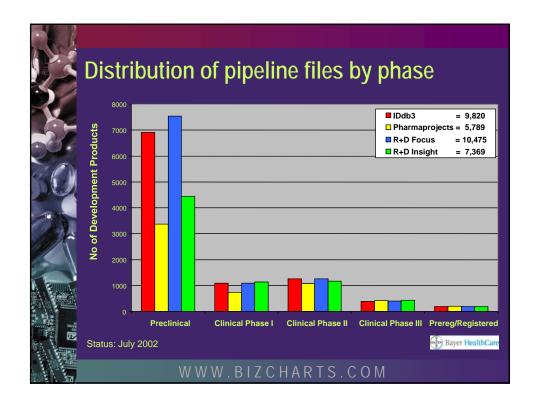
- Pharmaprojects 407 records
- R&D Focus 202 records
- R&D Insight 201 records
- IDdb 250 records
- Also searched Prous Drug Data Report -- 2880 records retrieved

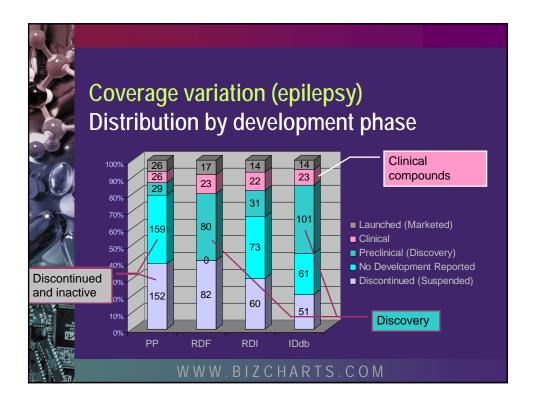


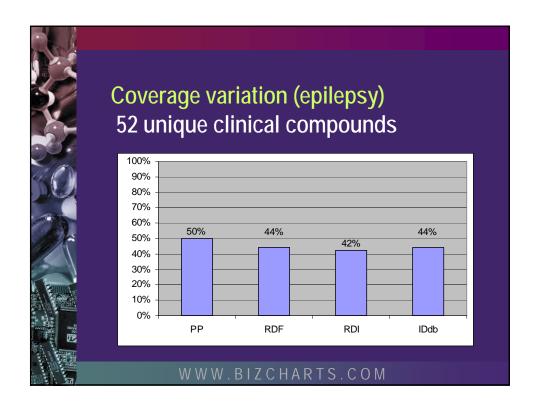


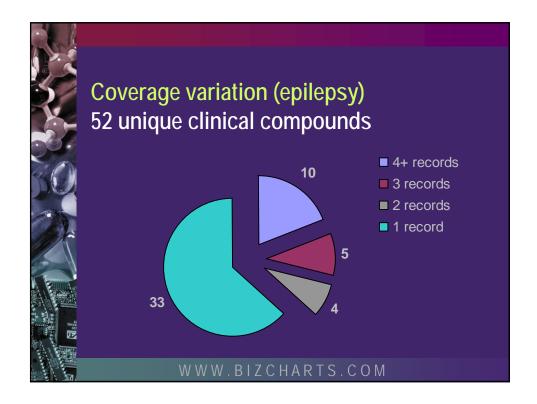


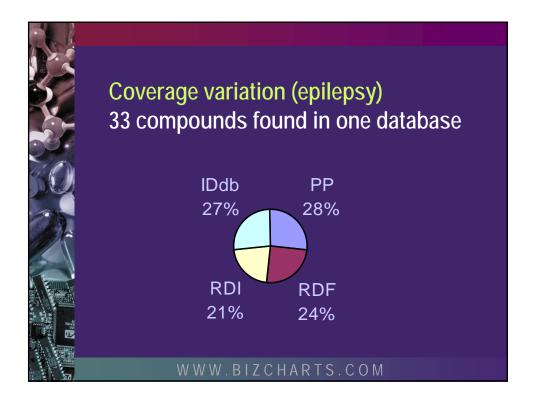


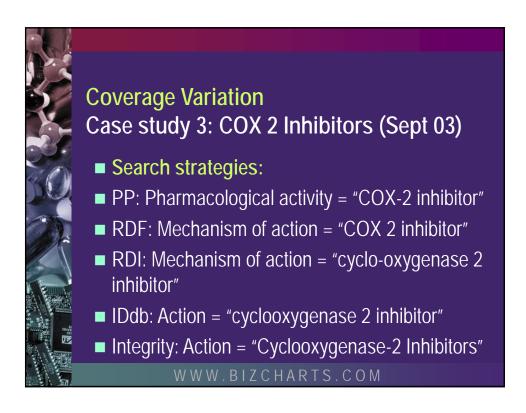


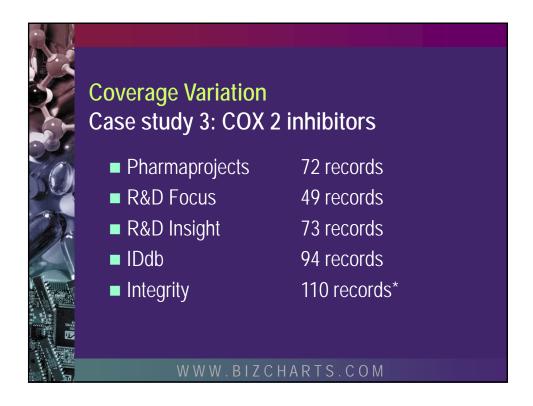


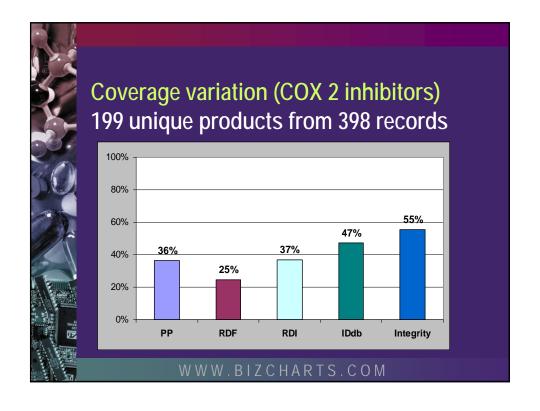














Coverage variation - conclusions
Only a fraction of total compounds will
be retrieved from any one database

- A search of any one database will retrieve about 40% of the unique compounds.
- The number of unique compounds will be roughly half the total number of records retrieved.

WWW BI7CHARTS COM



# Coverage variation - conclusions Over half the compounds will be retrieved from only one database

- Compounds (esp. in discovery) can be assigned *very* different names.
- Possible that similar compounds were not matched.
- Key reason: the search missed compounds due to indexing variations.



Coverage variation - conclusions
Variations in indexing result in much of
the coverage variation

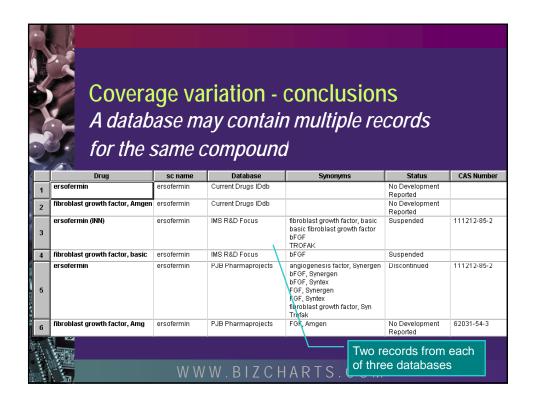
- Alfimeprase: doesn't appear in R&D Insight
- CNTF: RDI record doesn't mention Amgen
- Abelcet (ABLC): IDdb record mentions Amgen, but not indexed by Amgen
- KRN-568: Indexed as Kirin & NPS in Pharmaprojects

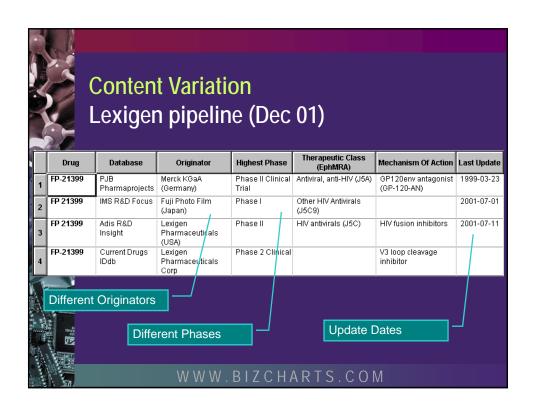
WWW.BIZCHARTS.COM

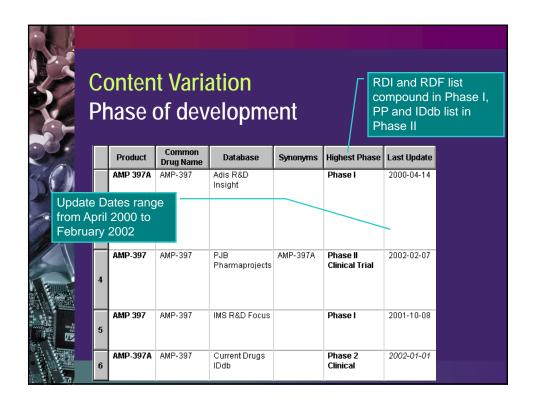


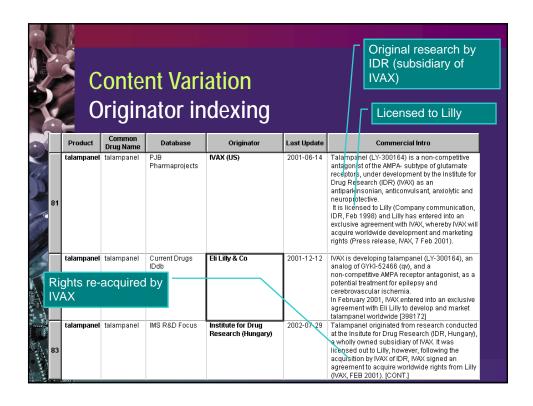
Coverage variation - conclusions Compounds are usually in all databases, but are not retrieved by the search

- Sanofi-Synthelabo analysis by Bayer
- Preclinical: of 30 records in IDdb, 15 were in the other 3 files.
- Phase I: 10 records in IDdb; all in other 3.
- Phase II: 12 records in IDdb; 10 in other 3.
- Phase III: 6 records in IDdb; all in other 3.











### Coverage and content Why is there so much variation?

- Drug pipeline records are not abstracted from a reference document (as with patents)
- Each record is an editorial view of the compound based on a range of sources.

WWW.BIZCHARTS.COM



## Coverage and content Information sources can be misleading

- "As a disclosure of compound information is balanced by the business need to maintain competitive advantage, some compound information has not been disclosed at this time."
  - -- AstraZeneca website(as reported by Alexander Mullin, Bayer)



### Coverage and content Why is there so much variation?

- Publishers differ in background and focus.
- Editors assigned by therapy or company -- vary in experience and skill.
- Different update schedules
- Different index vocabularies
- Unique content (e.g. Lehman data in RDI).

WWW.BIZCHARTS.COM

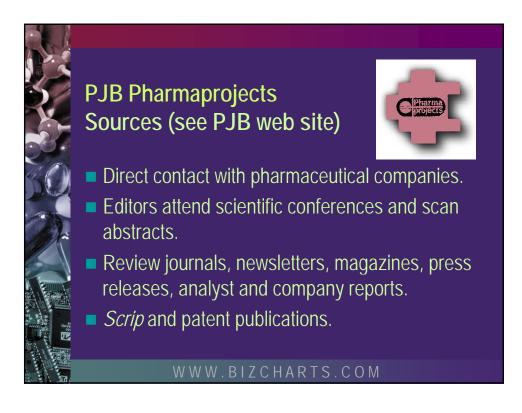


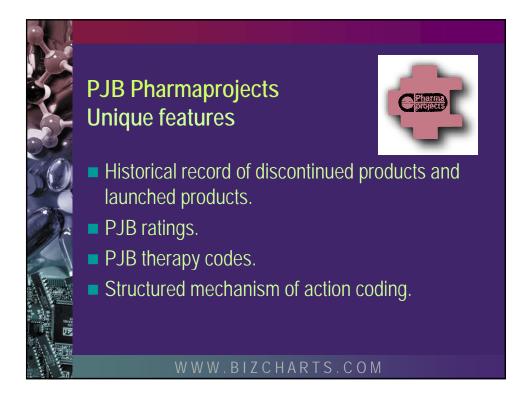
#### PJB Pharmaprojects Background

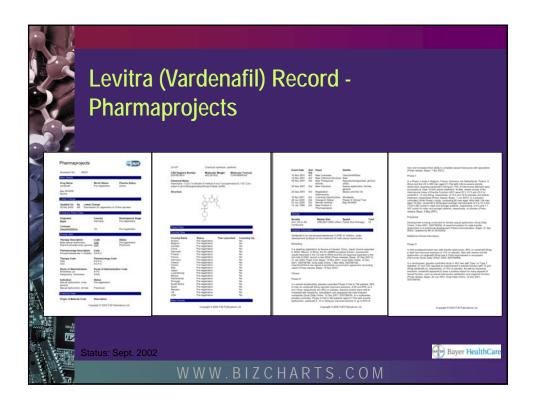


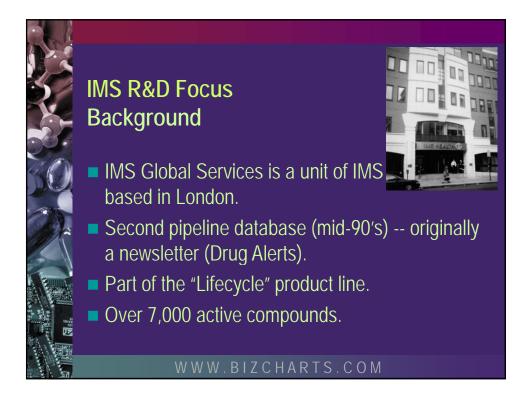
- PJB: named for founder Philip J. Brown; based in Richmond (outside of London)
- PJB's best known product is Scrip.
- The first pipeline database (1980?) -- originally published in books.
- Contains 26,000 compounds -- 6,000 active.

WWW.BIZCHARTS.COM

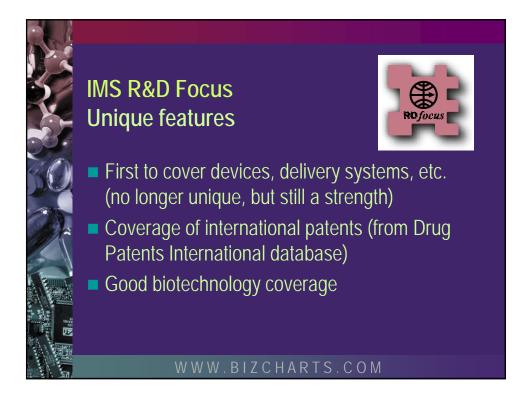




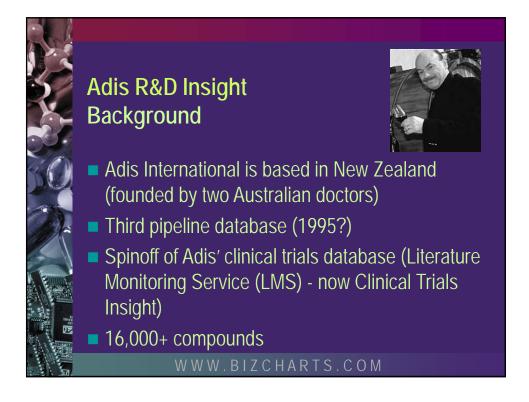


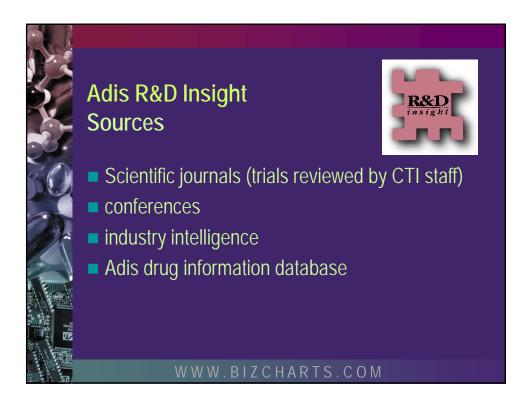


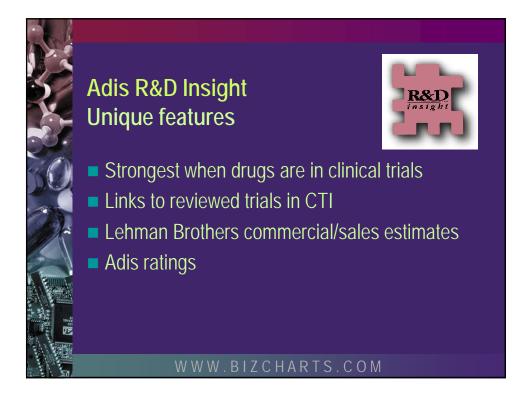




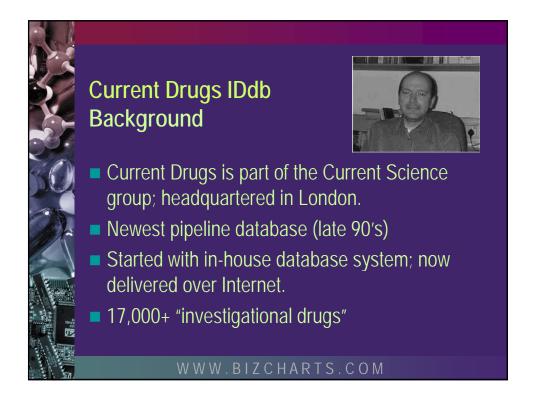


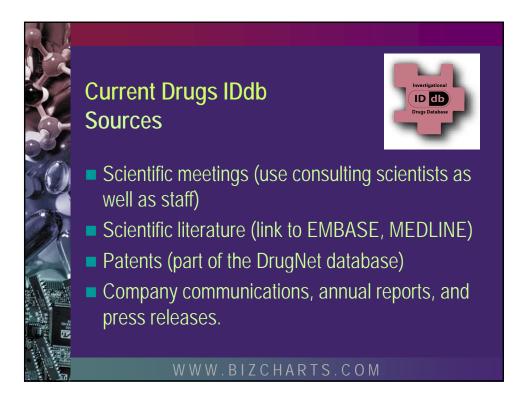


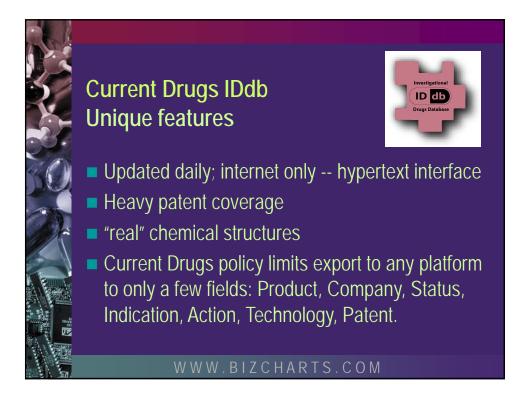


















### Surfing the drug pipeline Keep in mind...



- The good news: can leverage of work of dozens of researchers and editors.
- The bad news: must resolve the differences between these experts.
- Develop your own editorial policy for drug pipeline data!

WWW.BIZCHARTS.COM