Expression of laboratory examination results in medical literature

Takashi Okumura

National Institute of Public Health
taka@niph.go.jp
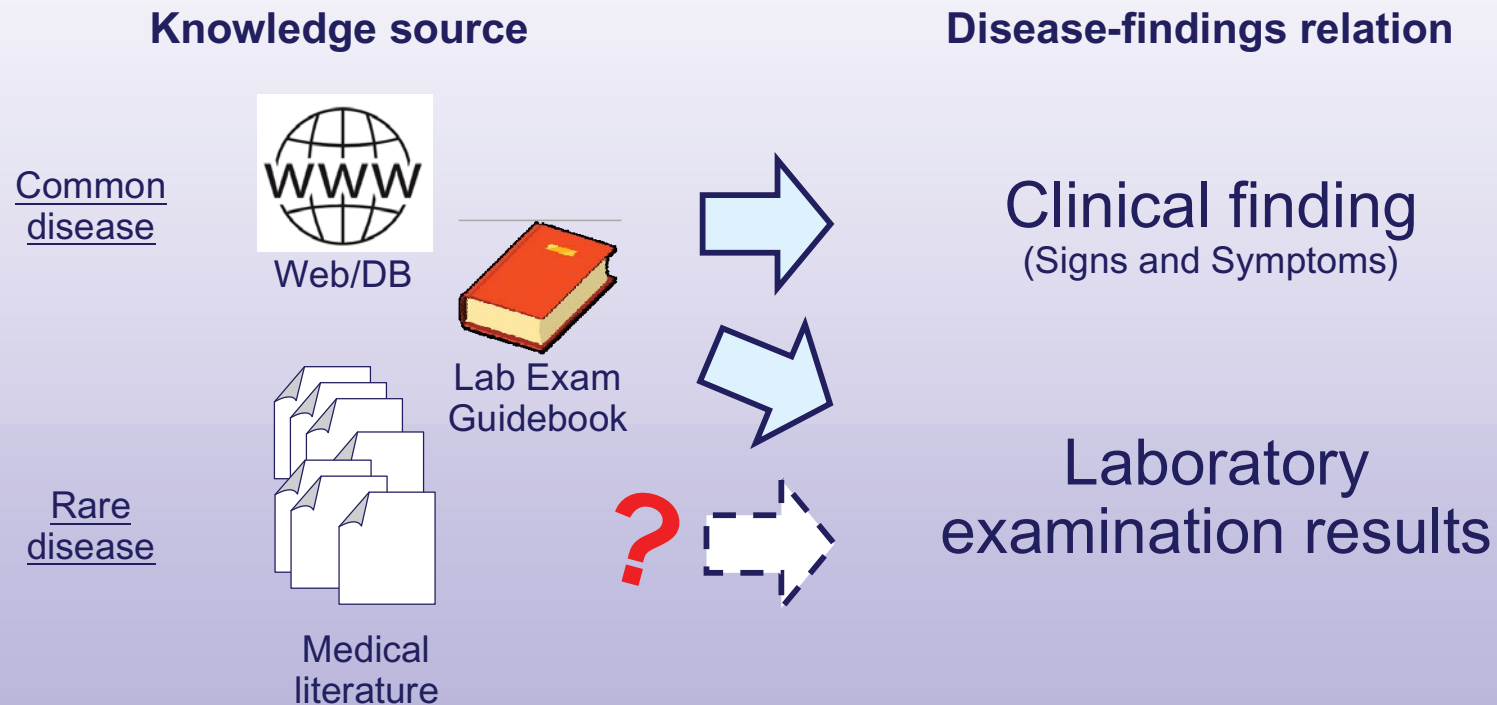
Eiji Aramaki
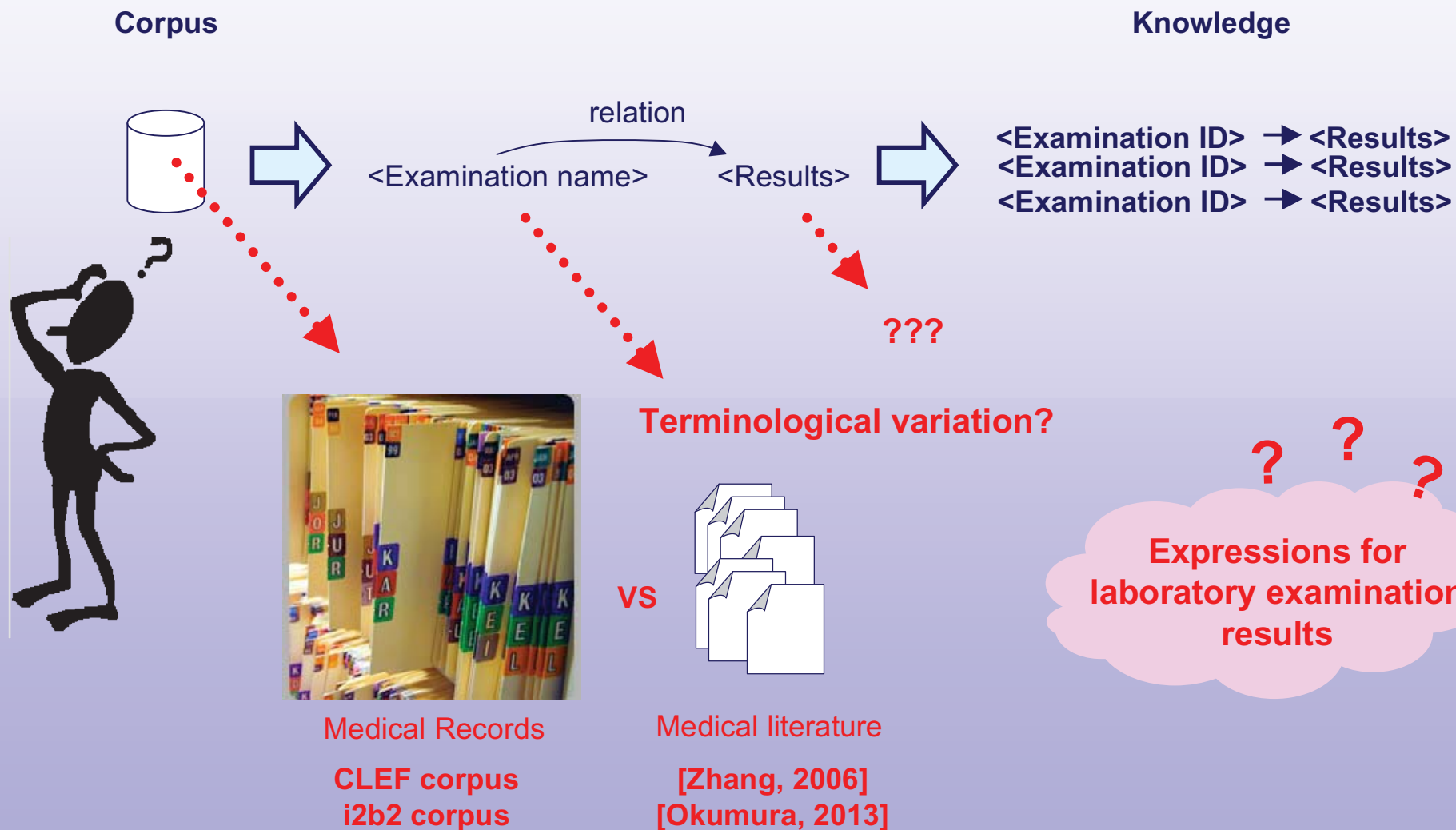
Kyoto University
eiji.aramaki@gmail.com

Yuka Tateisi *

National Institute of Informatics
yucca@nii.ac.jp

BioTxtM 2014

Disease Knowledge Base and NLP

Knowledge source

Disease-findings relation

Common
disease

Web/DB

Lab Exam
Guidebook

Clinical finding
(Signs and Symptoms)

Rare
disease

Medical
literature

?

**Laboratory
examination results**

Need to interpret expressions for laboratory examination results

Related Works

Corpus

Knowledge

relation

<Examination name>   <Results>

<Examination ID> ➡ <Results>
<Examination ID> ➡ <Results>
<Examination ID> ➡ <Results>

???

Terminological variation?

vs

Medical Records

Medical literature

**CLEF corpus
i2b2 corpus**

**[Zhang, 2006]
[Okumura, 2013]**
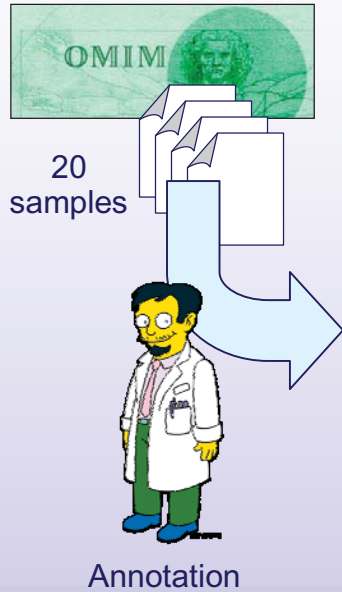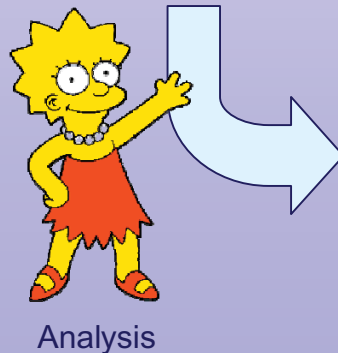
Expressions for laboratory examination results

Expressions for laboratory examination results

E1. elevated *mean pulmonary artery pressure*
E2. elevated *liver function enzymes*
E3. elevated *transaminases*
E4. moderately elevated *alkaline phosphatase*
E5. normal *esophageal manometry* findings
E6. *SGPT* was chronically elevated
E7. *hCG stimulation tests* showed normal testosterone responses

E8. deletion of part of the short arm of chromosome 5
E9. lymphocyte infiltration of the CNS
E10. low to low-normal gonadotropin (152760) resp
E11. Administration of testosterone enanthate was eff
E12. insensitive to growth hormone

E13. hypercholesterolemia

Examination name

+

Evaluative Expression

Evaluative Expression

20 samples

OMIM

Annotation

Analysis

		Evaluative Expression	
		+	-
Examination Name	+	Quantitative examinations	Qualitative examinations
	-	Descriptive expression	Nominal expression

(Simple) Experiment setting

Examination Dict

Evaluative Expression List

Nominal Result Dict

manual compilation

OMIM

orphanet

GENIA SS

GENIA SS

filter

filter

Count

Count
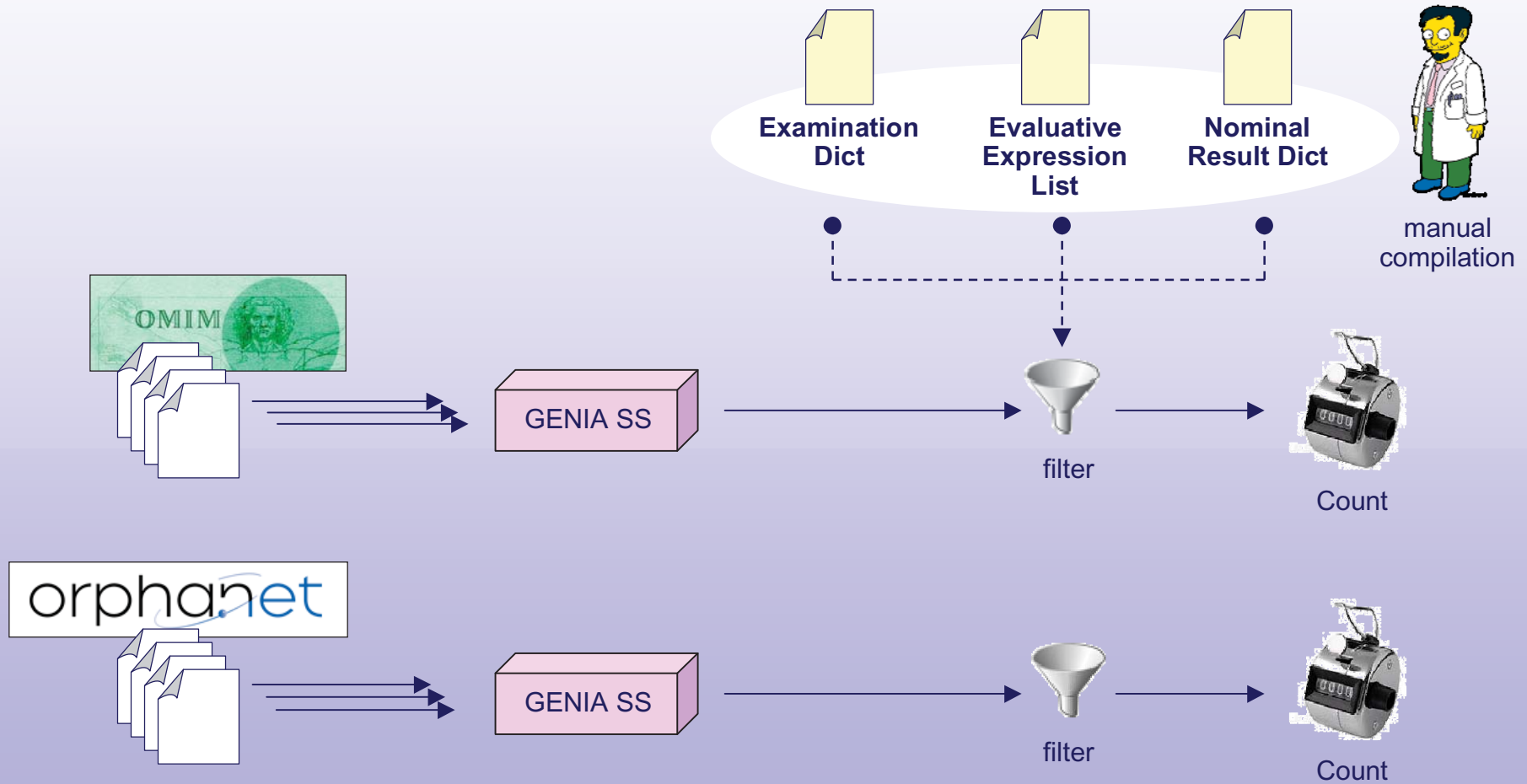
(Simple) Dictionary for examination names

Result Category	# of records	# of synonyms	
Numerical	334	532	(1.6)
High/Low	202	368	(1.8)
Positive/Negative	163	391	(2.4)
Load/Function test	29	52	(1.8)
Typing/Genetic test	21	32	(1.5)
Fraction/Isozymes	18	40	(2.2)
Microscopical findings	14	22	(1.6)
Others	37	60	(1.6)
Total	818	1497	(1.8)

→ **1497 items**

Fumimaro Takaku (ed.). 2009. *Laboratory examinations databook 2009-2010*. Igaku-Shoin.

Ex) Human leukocyte antigen typing → "HLA typing" and "HLA findings"

CSF → colony stimulating factor, cerebrospinal fluid
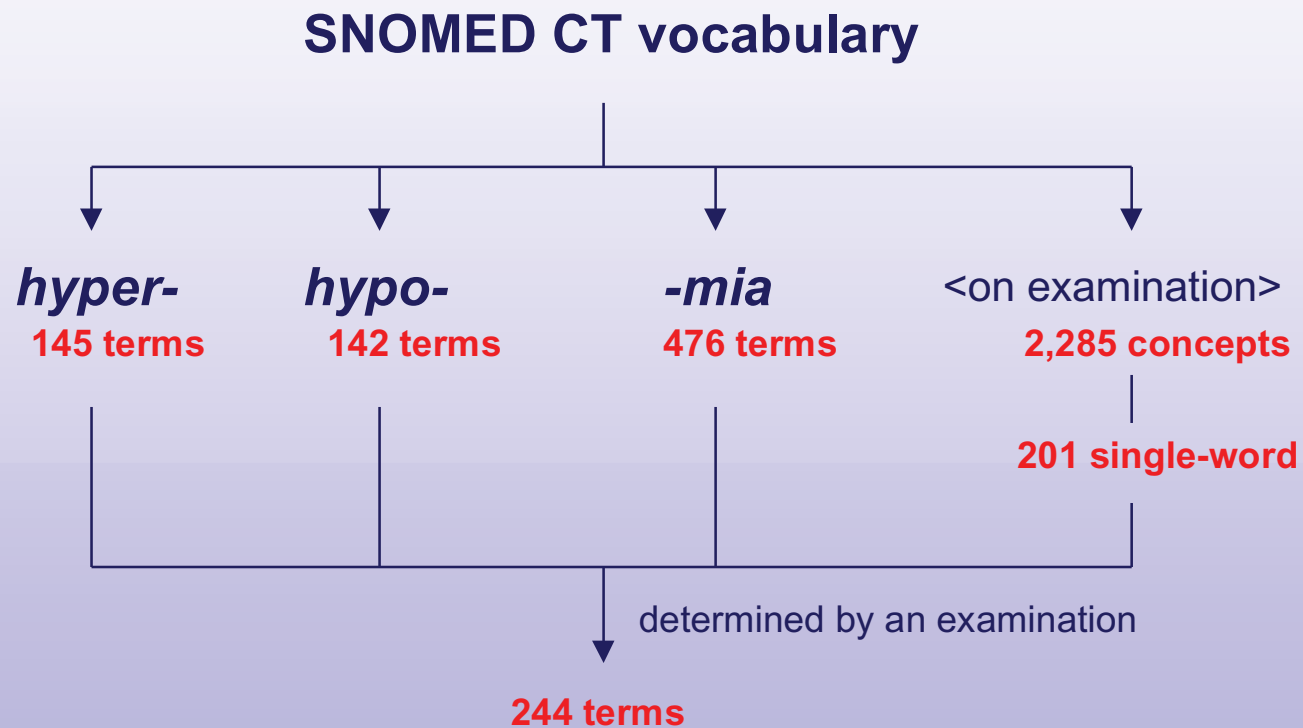
Creatine kinase isozymes → CK isozymes

demonstrated marked elevation of BB isozyme fraction of serum creatine kinase for male sibs with this disorder…!?

(Simple) Dictionary for evaluative expressions

Normal	Abnormal	Increase	Decrease	Positive	Negative	Sufficiency	Deficiency
normal	abnormal	increase	decrease	positive	negative	sufficiency	deficiency
sensitive	insensitive	increased	decreased	above	below	sufficient	deficient
effective	ineffective	increasing	decreasing	high	low	excess	insufficiency
	defective	elevated	declined	higher	lower	excesses	insufficient
		elevation	decline			excessive	
		elevating	declining				
		rise	lowering				
		risen	diminished				
		rising	diminishing				
			reduced				
3 items	4 items	9 items	10 items	4 items	4 items	5 items	4 items

43 items

Dictionary for nominal expressions for laboratory examination results

SNOMED CT vocabulary

hyper-	*hypo-*	*-mia*	<on examination>
145 terms	**142 terms**	**476 terms**	**2,285 concepts**

201 single-word

determined by an examination

244 terms

An expression of a finding or a medical state that is conclusively determined by a single laboratory examination, such as *hyperkalemia*.

(Simple) Experiment setting

1497 Examination Dict

43 Evaluative Expression List

244 Nominal Result Dict

manual compilation

OMIM

6,727 Diseases → GENIA SS → **587,601 Sentences** → filter → **193,687 Sentences** → filter → Count **?**

orphanet

6,442 Diseases → filter → **3,369 Diseases** → GENIA SS → **32,768 Sentences** → filter → Count **?**
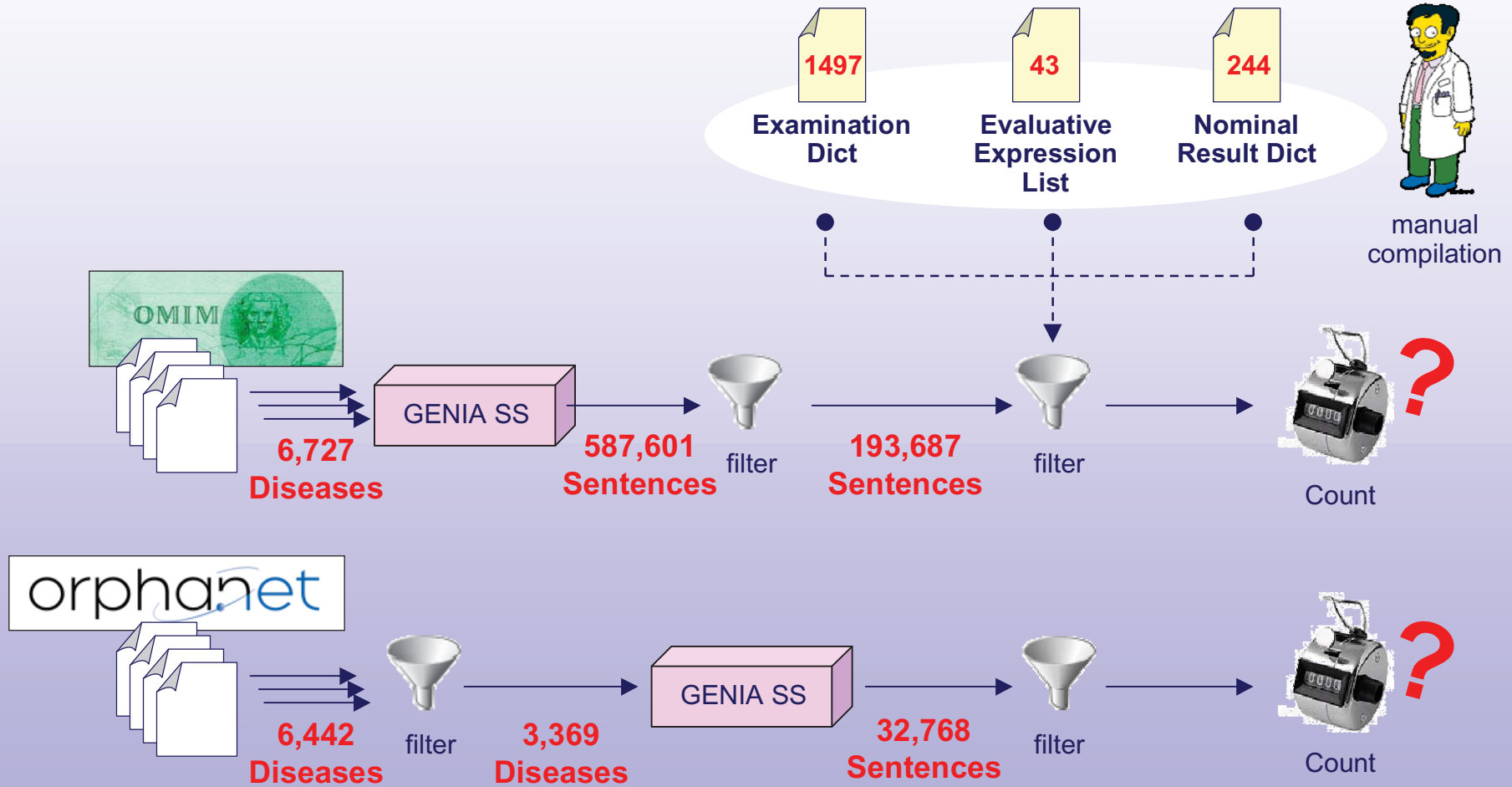
Result

	OMIM		Orphanet	
Sentences with Examination Name (EN+)	13,343	(6.9%)	2,555	(7.8%)
Sentences with Evaluative Expression (EE+)	30,577	(15.8%)	4,566	(13.9%)
Sentences with both EN+ and EE+	4,742	(2.4%)	755	(2.3%)
Sentences with nominal laboratory results	4,836	(2.5%)	1,165	(3.6%)
Total # of sentences	193,687	(100.0%)	32,768	(100.0%)

Table 5: Number of sentences with examination names and evaluative expressions
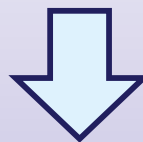
		Evaluative Expression		
		+	-	
Examination Name	+	2.3 −2.4%	4.5 −5.5%	6.9 −7.8%
	-	11.6 −13.4%	2.5 −3.6%	15.2 −15.9%
		13.9 −15.8%	7.0 −9.1%	

Analysis

		Evaluative Expression		
		+	-	
Examination Name	+	**2.3 − 2.4%**	4.5 − 5.5%	**6.9 − 7.8%**
	-	11.6 − 13.4%	**2.5 − 3.6%**	15.2 − 15.9%

above and *below* appear often in ordinary English, and adjectives such as *normal* and *abnormal* may be used also for clinical findings

Counted with limited vocabulary, and thus, the actual figure was underestimated

The number suggests undocumented classes of expressions that merit further investigation

Though the simple statistics is not sufficiently accurate to draw a firm conclusion...

Discussion

for future corpus design

- The newer corpus has to address the omission of examination names, in the expressions of laboratory examination results.

    Ex) "*fragility of red blood cells*" for a patient of paroxysmal nocturnal hemoglobinuria suggests *Ham test* (+)

- We identified eight categories for evaluative expressions, but degree of change must be taken into account

    Ex) "slight increase" vs "severe increase"

- Expressions of qualitative results also demand special handling

    Ex) Typing tests and genetic analysis

- Complexity in the expressions of examination results must be addressed

    Ex) because sentences displaying a simple relation between an examination name and an examination result are in the minority

Conclusion

- Expressions of laboratory examination results can be classified by existence of their examination names
  - *Domain experts omit formal names if there is more essential information than names in the context and if the description carries the information necessary to infer the names*

- Expressions of laboratory results tend to have evaluative expressions that can be used for another axis to classify
  - *The expressions are used for differentiating quantitative and qualitative results, each of which requires an independent strategy to interpret expression of laboratory examination results*

- This class of expressions is beyond the scope of the existing corpora designed for medical records, and further investigation is needed to interpret laboratory examination results in medical literature