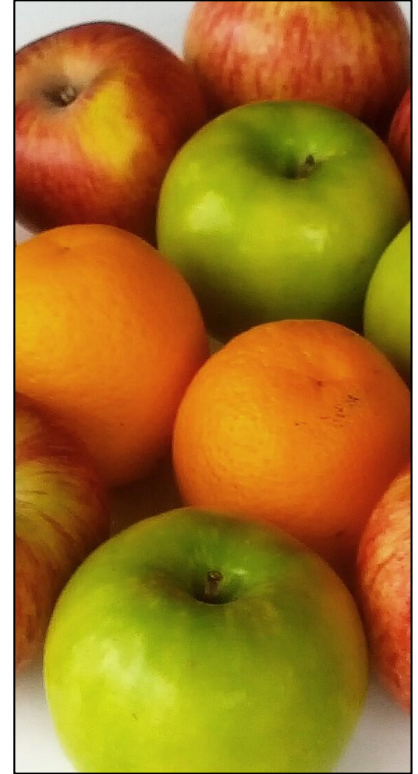


# Inter-coder agreement assessment for SNOMED CT use

Daniel Karlsson  
Linköping University, Sweden

# Outline

- Theory
  - What is (dis)agreement
  - How can agreement be assessed
  - What about agreement and SNOMED CT?
- Demonstrations throughout
- Questions



# Software used

- R
  - Open source statistics software with lots of packages
  - <https://www.r-project.org/>
- Kilem Gwet's implementation of agreement measures
  - [http://www.agreestat.com/r\\_functions.html](http://www.agreestat.com/r_functions.html)
- SNOMED CT specific function
  - <https://github.com/danka74/SnomedAgreement>



# (Dis)agreement

- Judgement variables, classification of phenomena into predefined classes
  - E.g. using SNOMED CT
- Two or more coders classifying the same phenomenon
- Agreement = the extent to which the results of the classifications coincide
- Disagreement detrimental to semantic interoperability

399210005 | neurological investigation (procedure) |  
268970009 | central nervous system examination (procedure) |  
271888005 | on examination - neurological (finding) |  
102957003 | neurological finding (finding) |  
271924005 | neurological test finding (observable entity) |  
271909001 | neurology examination finding (observable entity) |  
18373002 | nervous system function (observable entity) |  
...

# (Dis)agreement

- Judgement variables, classification of phenomena into predefined classes
  - E.g. using SNOMED CT
- Two or more coders classifying the same phenomenon
- Agreement = the extent to which the results of the classifications coincide
- **Disagreement detrimental to semantic interoperability**

399210005 | neurological investigation (procedure) |  
268970009 | central nervous system examination (procedure) |  
271888005 | on examination - neurological (finding) |  
102957003 | neurological finding (finding) |  
271924005 | neurological test finding (observable entity) |  
271909001 | neurology examination finding (observable entity) |  
18373002 | nervous system function (observable entity) |  
...

# (Dis)agreement data

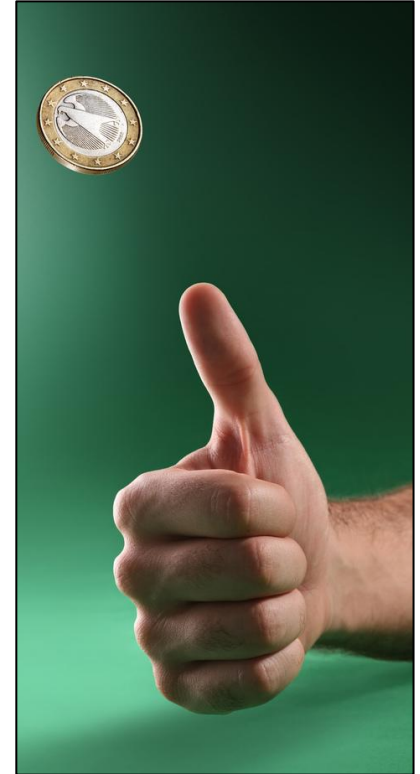
	User 1	User 2	User 3	User 4	User 5	User 6
Case 1	8517006   Ex-smoker (finding)	702975009   Ex-tobacco user (finding)	8517006   Ex-smoker (finding)	8517006   Ex-smoker (finding)	8517006   Ex-smoker (finding)	8517006   Ex-smoker (finding)
Case 2	229819007   Tobacco use and exposure (observable entity)	229819007   Tobacco use and exposure (observable entity)	266918002   Tobacco smoking consumption (observable entity)	365981007   Finding of tobacco smoking behavior (finding) 	229819007   Tobacco use and exposure (observable entity)	365981007   Finding of tobacco smoking behavior (finding) 
Case 3	77176002   Smoker (finding)	77176002   Smoker (finding)	77176002   Smoker (finding)	77176002   Smoker (finding)	77176002   Smoker (finding)	77176002   Smoker (finding)
...						

# Chance (dis)agreement

- Two persons flip a coin, probability that they will “agree” is 50 % (percentage agreement)

	Head	Tail
Head	2470	2477
Tail	2513	2540

$$( 2470 + 2540 ) / 10000 = 0.501$$



# Assessment of agreement

- Agreement measures
  - Percentage agreement
  - For example Cohen's Kappa
- 0 means no agreement
- 1 means perfect agreement





# Scale of agreement

- Nominal scale agreement
  - Coders either fully agree or fully disagree
- Ordinal scale, weighted agreement
  - Semantic distance between concepts

75367002 | blood pressure  
(observable entity) |

251076008 | non-invasive  
arterial pressure  
(observable entity) |

392570002 | blood  
pressure finding (finding) |

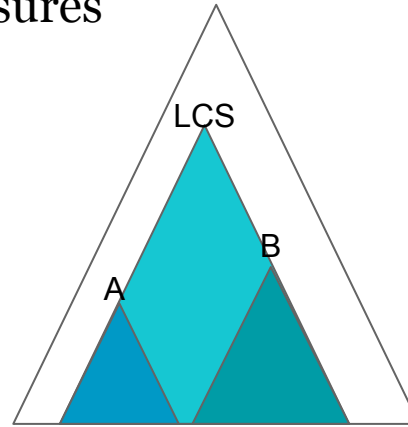
6973005 | blood pressure  
taking (procedure) |

371911009 | measurement  
of blood pressure using  
cuff method (procedure) |

...

# Semantic distance measures

- Use-case dependant
- Path-length based measures
  - Path length to least common subsumer
- Information-content based measures
  - E.g. Lin



75367002 | blood pressure  
(observable entity) |

251076008 | non-invasive  
arterial pressure  
(observable entity) |

392570002 | blood  
pressure finding (finding) |

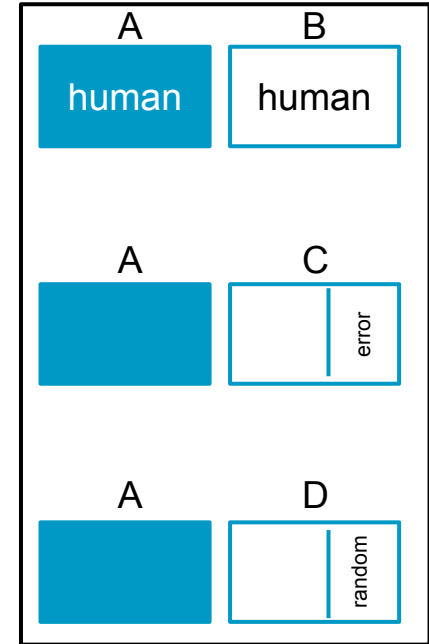
6973005 | blood pressure  
taking (procedure) |

371911009 | measurement  
of blood pressure using  
cuff method (procedure) |

...

# Example 1: National procedure classification mapping

- Two human coders A and B
  - 490 procedure codes cross-mapped from NCSP to SNOMED CT
- Percentage agreement 72 %
- Path length to LCS
- Datasets
  - Dataset AB: human coders
  - Dataset AC: coder A + “novice” coding errors introduced
  - Dataset AD: coder A + random codes



# Agreement with semantic distance

	<b>Dataset AB (human-human)</b>	<b>Dataset AC (human-novice)</b>	<b>Dataset AD (human-random)</b>
<b>Nominal <math>\alpha</math>, mean (95 % CI)</b>	0.72 (0.68-0.76)	0.72 (0.68-0.76)	0.72 (0.68-0.76)
<b>Semantic <math>\alpha</math>, mean (95 % CI)</b>	0.89 (0.86-0.92)	0.84 (0.80-0.88)	0.47 (0.41-0.53)

## Example 2: Terminology binding

- ASSESS CT project
- Binding to information model extracts
  - 12 extracts, 101 elements in total
  - 6 coders from 6 countries
- Comparing SNOMED CT to alternative terminologies

The screenshot displays a software interface for terminology binding. On the left, a sidebar lists various medical categories such as 'Cases', 'Adverse Reaction', and 'Smoking status'. The main area shows a 'Description' section with a 'HEART FAILURE CLINIC FIRST VISIT SUMMARY' and a 'Risk Factors' section containing a 'Tobacco Use Summary'. A dropdown menu for 'Smoking Status' is open, showing options: 'Current Smoker', 'Quitting', 'Ex-smoker', and 'Never Smoked'. Below this, an 'Information model attribute' table lists 'Smoking Status' with a value set of 'Precond. SNC' and 3 partial covers. A 'Value set' table below that lists the four smoking status categories with their respective SNOMED codes and coverage details.

Information model attribute	Value set
Smoking Status	Precond. SNC   3 Partial cover.   229619007 Tobacco use and ex
Overall meaning of value set	Precond. SNC   3 Partial cover.   365980008 Finding of tobacco u
Current Smoker	Precond. SNC   1 Full coverage.   77179002 Smoker (finding)
Quitting	Precond. SNC   1 Full coverage.   160616005 Trying to give up sm
Ex-smoker	Precond. SNC   1 Full coverage.   8517006 Ex-smoker (finding)
Never Smoked	Precond. SNC   1 Full coverage.   266819005 Never smoked tobac

# Acknowledgements

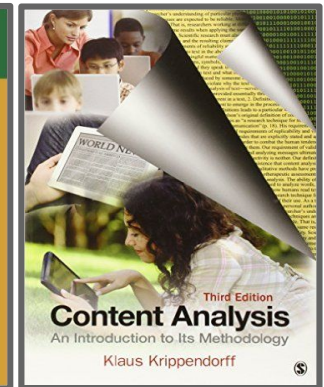
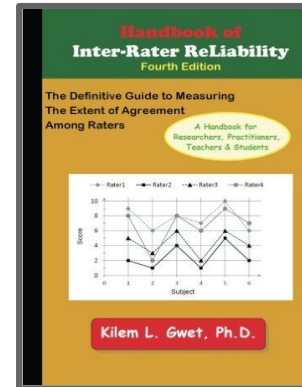
Kirstine Rosenbeck-Gøeg

Anne Randorff Højen

Håkan Örman

# References

Karlsson, D., Gøeg, K. R., Örman, H., & Højen, A. R. (2014, September). Semantic Krippendorff's  $\alpha$  for measuring inter-rater agreement in SNOMED CT coding studies. In MIE (pp. 151-155).



Thanks! Questions?

[www.liu.se](http://www.liu.se)



# Running R with ASSESS CT data

```
> krippen.alpha.raw(sct)
```

```
Krippendorff's Alpha Coefficient
```

```
=====
```

```
Percent agreement: 0.6620092 Percent chance agreement: 0.008904917
```

```
Krippendorff alpha coefficient: 0.6589724 Standard error: 0.03617237
```

```
95 % Confidence Interval: ( 0.5871895 , 0.7307553 )
```

```
P-value: 0
```

```
> krippen.alpha.raw(sct, weight)
```

```
Krippendorff's Alpha Coefficient
```

```
=====
```

```
Percent agreement: 0.782077 Percent chance agreement: 0.07926808
```

```
Krippendorff alpha coefficient: 0.7633155 Standard error: 0.02885617
```

```
95 % Confidence Interval: ( 0.7060514 , 0.8205796 )
```

```
P-value: 0
```