# Comparing high-scoring and low-scoring REF2014 Impact Case Studies

Bella Reichard, Newcastle University

- b.reichard2@ncl.ac.uk, Twitter: @bellareichard

Birmingham Summer School, 27th June 2018

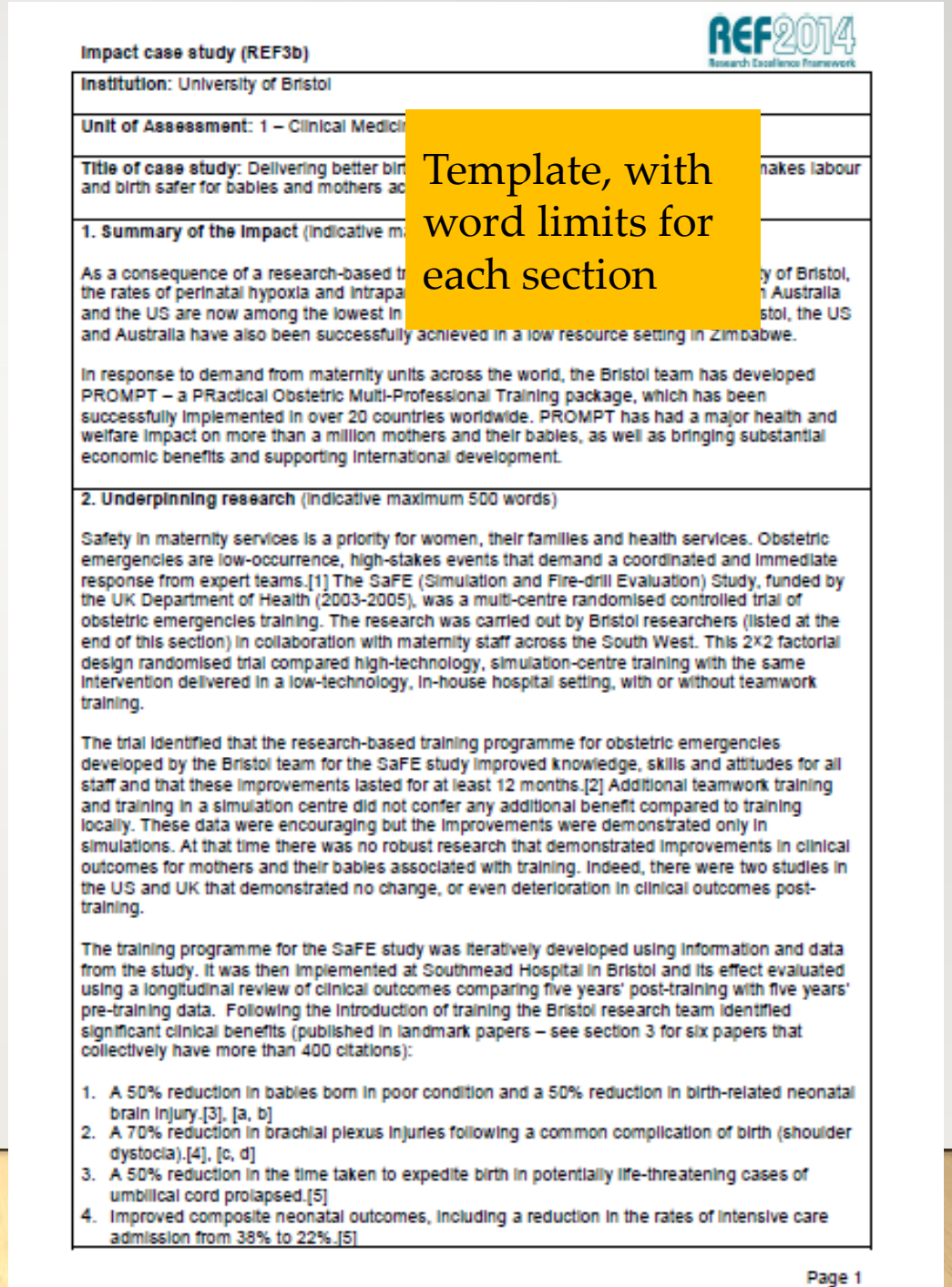
# Context

- UK Funding council "Research Excellence Framework" (REF)
- 2014: included "Impact Case Studies" (6,679 submitted)

- Impact: "**an effect** on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, **beyond academia"** (REF, 2011: 48)

# Structure of Impact Case Study

- Section 1: Summary of the impact
- Section 2: Underpinning research
- Section 3: References to the research
- Section 4: Details of the impact
- Section 5: Sources to corroborate the impact
- My corpus: text only (Sections 1,2,4)



Impact case study (REF3b)

REF2014 Research Excellence Framework

**Institution:** University of Bristol

**Unit of Assessment:** 1 – Clinical Medici

**Title of case study:** Delivering better birt … makes labour and birth safer for babies and mothers ac

Template, with word limits for each section

**1. Summary of the impact** (indicative m…

As a consequence of a research-based t… …y of Bristol, the rates of perinatal hypoxia and intrapa… …n Australia and the US are now among the lowest in … …stol, the US and Australia have also been successfully achieved in a low resource setting in Zimbabwe.

In response to demand from maternity units across the world, the Bristol team has developed PROMPT – a PRactical Obstetric Multi-Professional Training package, which has been successfully implemented in over 20 countries worldwide. PROMPT has had a major health and welfare impact on more than a million mothers and their babies, as well as bringing substantial economic benefits and supporting international development.

**2. Underpinning research** (indicative maximum 500 words)

Safety in maternity services is a priority for women, their families and health services. Obstetric emergencies are low-occurrence, high-stakes events that demand a coordinated and immediate response from expert teams.[1] The SaFE (Simulation and Fire-drill Evaluation) Study, funded by the UK Department of Health (2003-2005), was a multi-centre randomised controlled trial of obstetric emergencies training. The research was carried out by Bristol researchers (listed at the end of this section) in collaboration with maternity staff across the South West. This 2×2 factorial design randomised trial compared high-technology, simulation-centre training with the same intervention delivered in a low-technology, in-house hospital setting, with or without teamwork training.

The trial identified that the research-based training programme for obstetric emergencies developed by the Bristol team for the SaFE study improved knowledge, skills and attitudes for all staff and that these improvements lasted for at least 12 months.[2] Additional teamwork training and training in a simulation centre did not confer any additional benefit compared to training locally. These data were encouraging but the improvements were demonstrated only in simulations. At that time there was no robust research that demonstrated improvements in clinical outcomes for mothers and their babies associated with training. Indeed, there were two studies in the US and UK that demonstrated no change, or even deterioration in clinical outcomes post-training.

The training programme for the SaFE study was iteratively developed using information and data from the study. It was then implemented at Southmead Hospital in Bristol and its effect evaluated using a longitudinal review of clinical outcomes comparing five years' post-training with five years' pre-training data. Following the introduction of training the Bristol research team identified significant clinical benefits (published in landmark papers – see section 3 for six papers that collectively have more than 400 citations):

1. A 50% reduction in babies born in poor condition and a 50% reduction in birth-related neonatal brain injury.[3], [a, b]
2. A 70% reduction in brachial plexus injuries following a common complication of birth (shoulder dystocia).[4], [c, d]
3. A 50% reduction in the time taken to expedite birth in potentially life-threatening cases of umbilical cord prolapsed.[5]
4. Improved composite neonatal outcomes, including a reduction in the rates of intensive care admission from 38% to 22%.[5]

Page 1

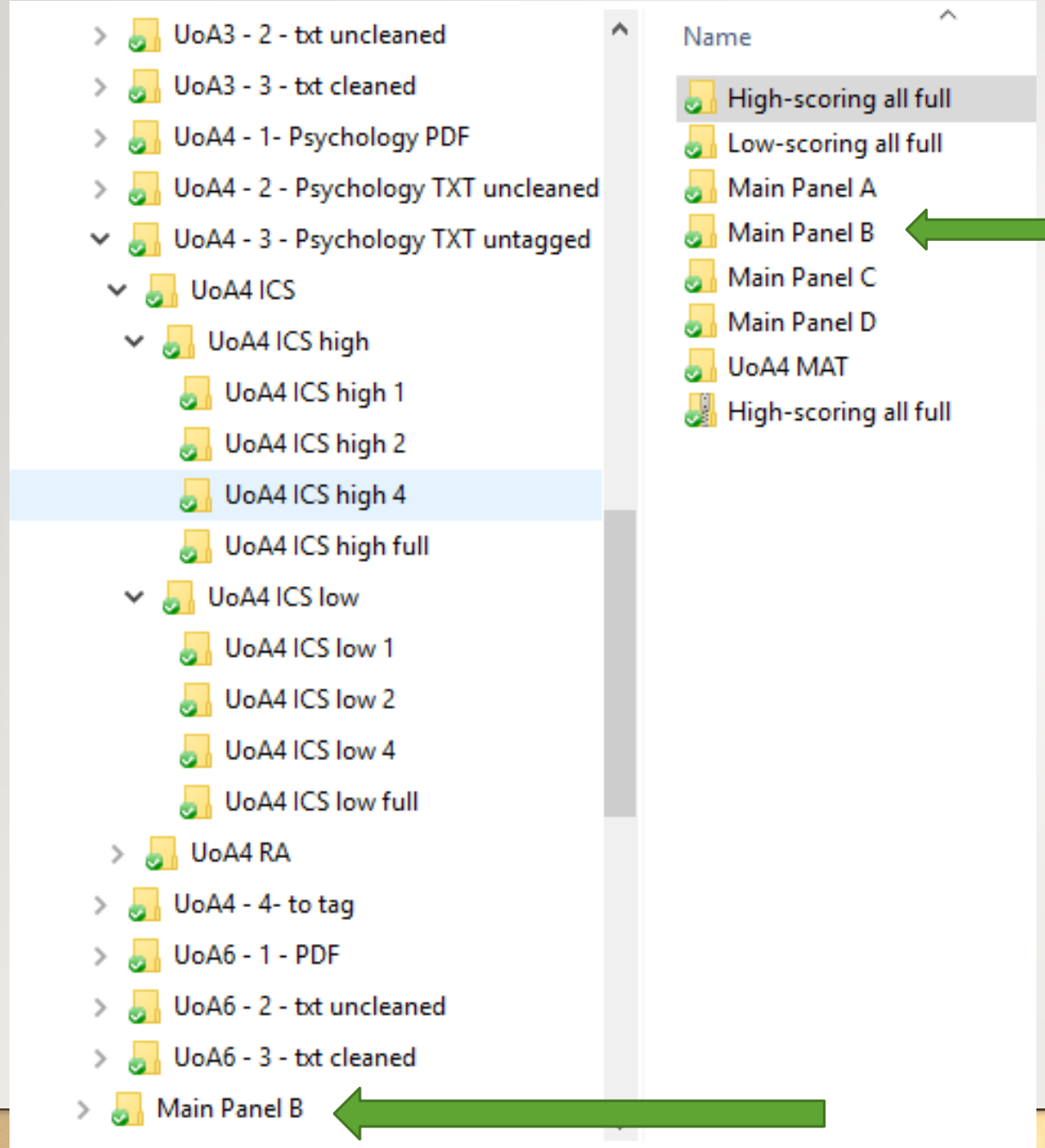
# Broad Research Questions

1. Are there differences between Impact Case Studies that received high and those that were given low scores?
2. Does this differ across “disciplines”? (36 Units of Assessments, broad discipline areas)

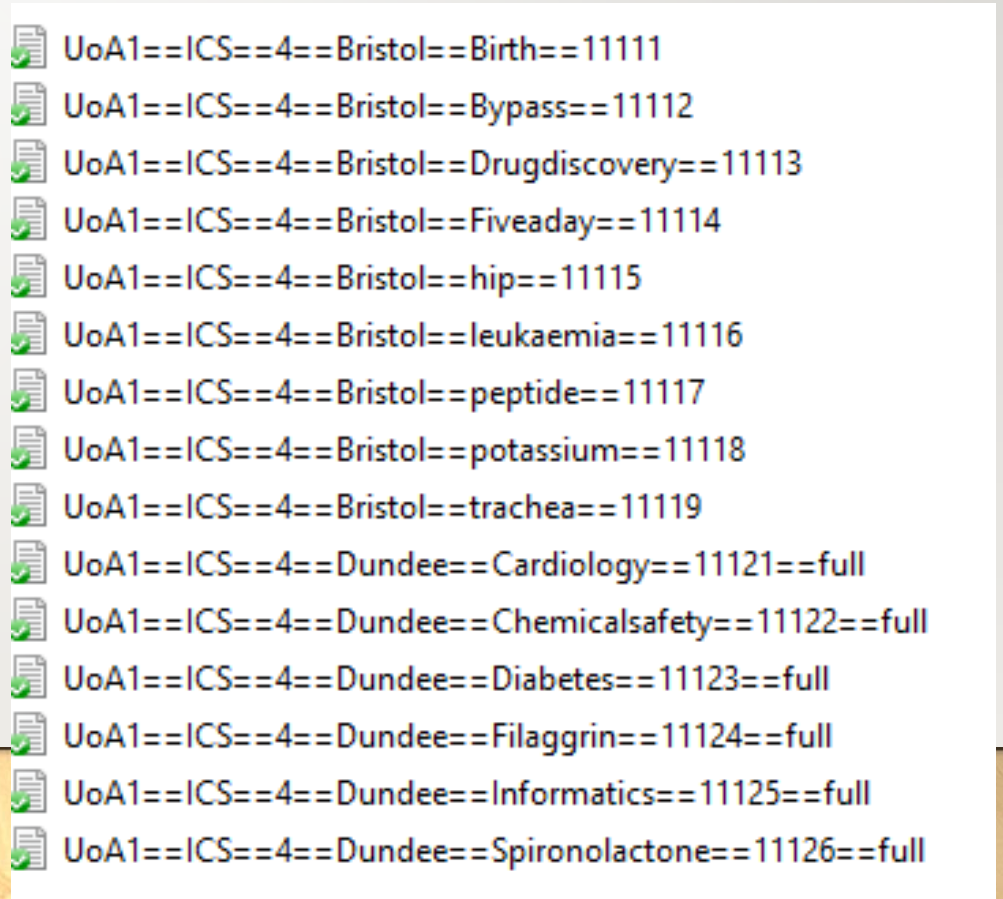
# The Corpus – Sample

| Main Panel | High-scoring (4*) | Low-scoring (1*/2*) | Total |
|---|---|---|---|
| A/B (Science) | 50 | 14 | 64 |
| C (Social Science) | 37 | 53 | 90 |
| D (Humanities) | 37 | 26 | 63 |
| **total** | **124** | **93** | **217** |

# Corpus files structure




Main Panel
-> Unit of Assessment
-> file type (PDF / converted / cleaned)
-> high- or low-scoring (subcorpus)
-> section of the text



# The Corpus – Size

| | High-scoring | Low-scoring | Total |
|---|---|---|---|
| **Files** | 124 | 93 | 217 |
| **Tokens (Lancsbox)** | Ca. 220,000 | Ca. 131,000 | Ca. 350,00 |
| **Types (Lancsbox)** | Ca. 17,000 | Ca. 11,000 | |
| **Type-Token-Ratio?** | | | |

# The Corpus – Preparation

- Download PDFs from REF website
- Use AntFileConverter to produce txt files -> no HTML to be preserved…
- Clean the txt files:
  - Delete template words (“indicative word limit: 100”, university name, page numbers…)
  - Delete “referencing” sections (sections 3 and 5)
  - Delete figures/tables but keep captions
- No further annotation or tagging (though Lancsbox adds lemmatisation and POS tagging) **– is it worth it for my research question?**

# Searches and tools used – in pilot study (one Unit of Assessment, out of 20)

- Wordlists (AntConc)
- Keyword searches (Lancsbox and AntConc)
- 3-grams and 4-grams (AntConc)
- KWIC for Keywords and 3/4-gram (Lancsbox -> Whelk)
- Collocations for Keywords and 3/4-grams (Lancsbox -> GraphColl)

# Searches – planned / desired

- **Same as for pilot study** (KW, n-grams, qualitative KWIC and collocates of these)
- **Qualitative analysis of KWIC lines** (principled collection of search terms, tbc)
  - Metadiscourse – stance and engagement markers?
  - Certain types of verbs?
- **Grammatical characteristics,** e.g. do low-scoring case studies use more passives, more long noun phrases…?
- **Cluster analysis**?
  - To see if there are (register) clusters corresponding to the sub corpora (high/low, certain Units of Assessment)?

Maybe a cluster analysis can help decide which words/features to examine more closely?

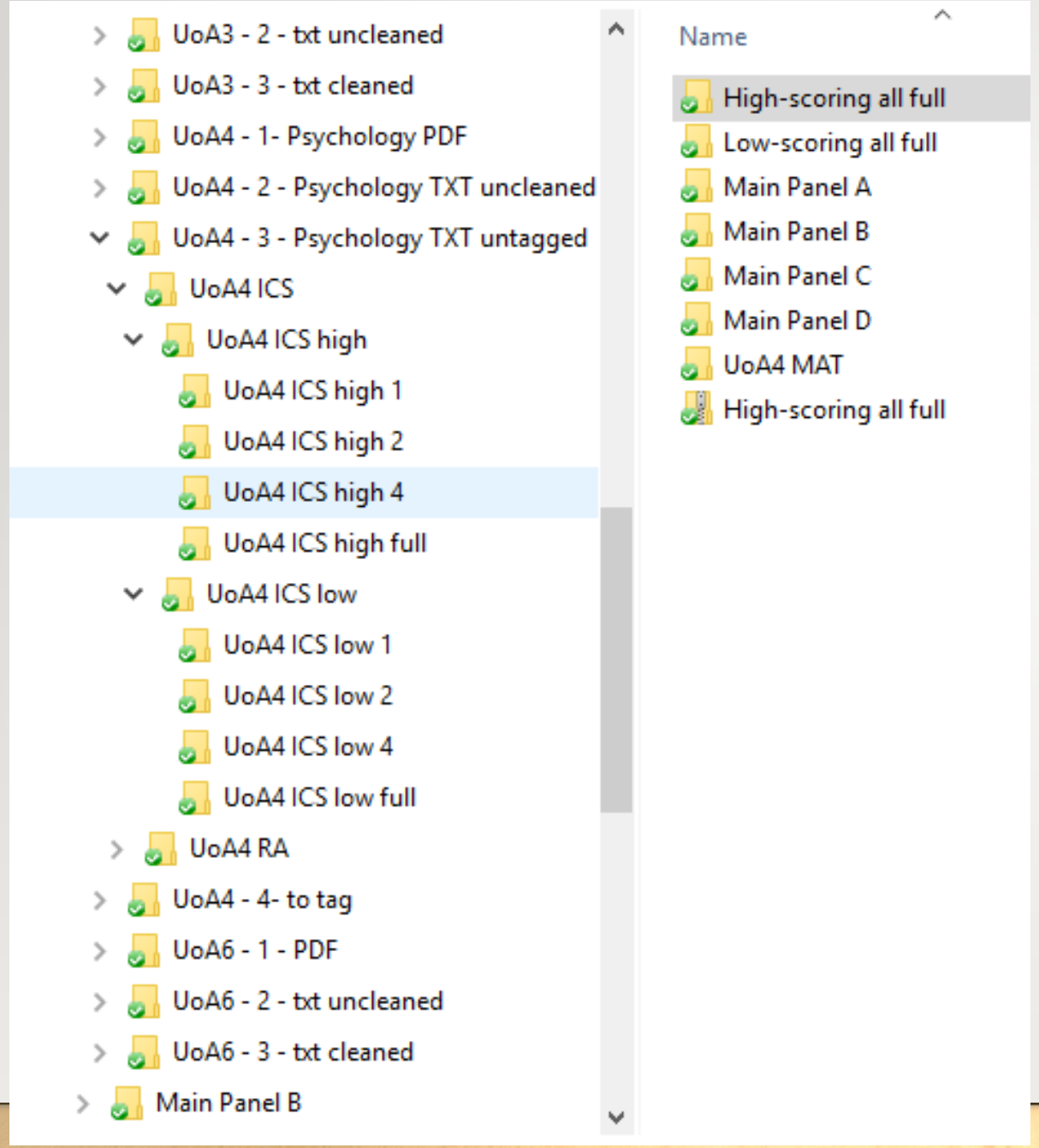
# My questions – corpus preparation

- **Balance** between time spent on corpus creation and analysis?

  -> specialist corpus for **my own use** to answer a genre question, rather than to describe language system

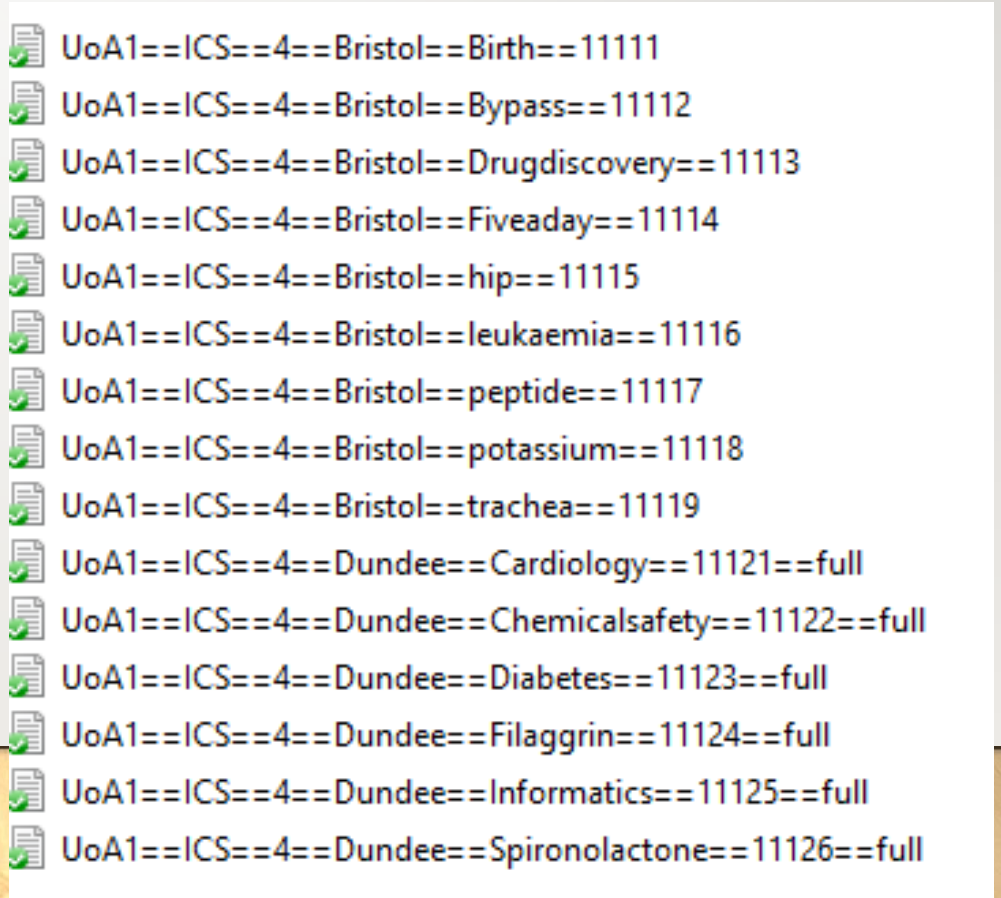
  -> more focus on (frequency-based or manual) analysis than on e.g. tagging?

- Currently don't have headers – do I need them if clue is in file name?




# Corpus files structure

Main Panel
-> Unit of Assessment
-> file type (PDF / converted / cleaned)
-> high- or low-scoring (subcorpus)
-> section of the text



# My questions – corpus preparation

- **Balance** between corpus creation and analysis? -> specialist corpus for **my own use** to answer a genre question, rather than to describe language system -> more focus on (frequency-based or manual) analysis than on e.g. tagging?
- Currently don't have headers – do I need them if clue is in file name?
- How do I deal with different dimensions/comparisons, e.g. high vs low across different disciplinary areas? Do I need headers for this?

# My questions – corpus analysis

- What statistical analysis can I do with a corpus of this **size** (just over **350,000** tokens) and for my **research questions** ("are there linguistic differences…")?
  - Cluster analysis – or is it too small?
  - Logistic regression – depends on the number of independent variables and the levels? (Independent variables = score and UoA = only 2? No need for regression? Or is score a dependent variable and my language features are explanatory factors?)
  - …or forget R and stick with Keyword and Collocation measures provided by AntConc etc.?
- How do I get there – from plain text files via annotated/tagged text files to a CSV (?) that is a dataset to be read in R for one of the analyses above…?

# My questions – how-to

- How do I avoid skewing – which measure to use for Keywords?
  - e.g. “education” appears as top 10 keyword in the low-scoring corpus, but this is because 23% case studies are from “education” (vs. 6%, in high-scoring)
- How do I code concordance lines for qualitative analysis? Export and do in Excel?
  - E.g. if I want to exclude examples from quotations – or ideally, count and analyse these separately

# What next?

- Any other ideas what to do with my corpus and how to get to interesting answers?

- b.reichard2@ncl.ac.uk, Twitter: @bellareichard