Fundamental Statistical Concepts

BIOS 6611

CU Anschutz

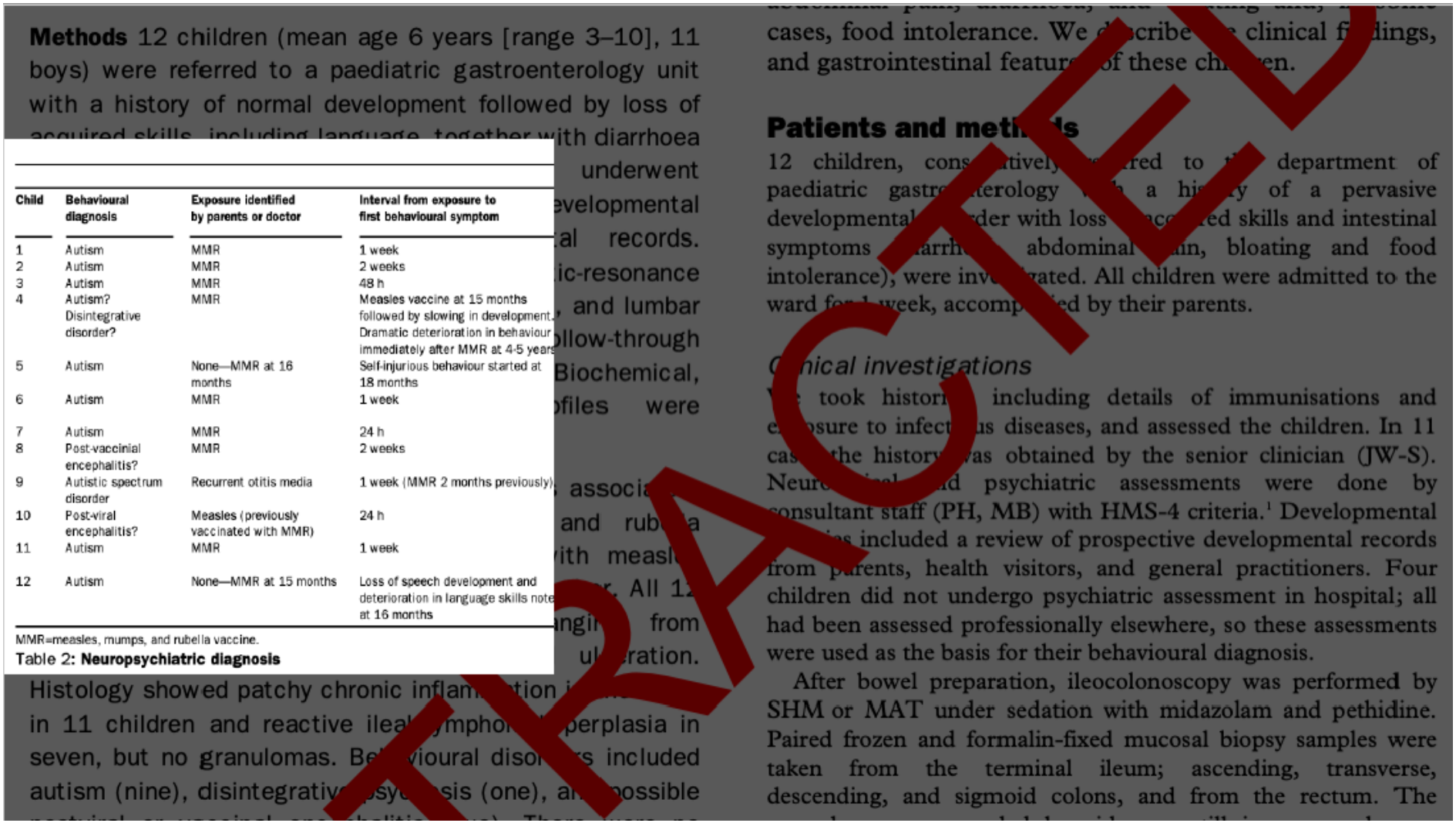
Week 1

1. What is statistics/biostatistics?
2. Populations and Samples

What is statistics/biostatistics?

Statistics/Biostatistics

- The use of data to answer scientific questions
- The discipline of *quantifying uncertainty*
  - Humans are really bad at this. Often our brains see patterns where there are none!
  - Statistics helps us distinguish between pure chance and meaningful patterns



Methods 12 children (mean age 6 years [range 3–10], 11 boys) were referred to a paediatric gastroenterology unit with a history of normal development followed by loss of acquired skills, including language, together with diarrhoea

Child	Behavioural diagnosis	Exposure identified by parents or doctor	Interval from exposure to first behavioural symptom
1	Autism	MMR	1 week
2	Autism	MMR	2 weeks
3	Autism	MMR	48 h
4	Autism? Disintegrative disorder?	MMR	Measles vaccine at 15 months followed by slowing in development. Dramatic deterioration in behaviour immediately after MMR at 4-5 years
5	Autism	None—MMR at 16 months	Self-injurious behaviour started at 18 months
6	Autism	MMR	1 week
7	Autism	MMR	24 h
8	Post-vaccinial encephalitis?	MMR	2 weeks
9	Autistic spectrum disorder	Recurrent otitis media	1 week (MMR 2 months previously)
10	Post-viral encephalitis?	Measles (previously vaccinated with MMR)	24 h
11	Autism	MMR	1 week
12	Autism	None—MMR at 15 months	Loss of speech development and deterioration in language skills noted at 16 months

MMR=measles, mumps, and rubella vaccine.

Table 2: **Neuropsychiatric diagnosis**

underwent developmental records. ic-resonance , and lumbar llow-through Biochemical, files were associa and rub a ith measl . All 12 angi from ul ration.

Histology showed patchy chronic inflam tion i in 11 children and reactive ileal lympho perplasia in seven, but no granulomas. Be vioural diso rs included autism (nine), disintegrativ psyc sis (one), a possible

cases, food intolerance. We describe the clinical findings, and gastrointestinal features of these children.

Patients and methods

12 children, consecutively referred to the department of paediatric gastroenterology with a history of a pervasive developmental disorder with loss of acquired skills and intestinal symptoms (diarrhoea, abdominal pain, bloating and food intolerance), were investigated. All children were admitted to the ward for 1 week, accompanied by their parents.

Clinical investigations

We took histories including details of immunisations and exposure to infectious diseases, and assessed the children. In 11 cases the history was obtained by the senior clinician (JW-S). Neurological and psychiatric assessments were done by consultant staff (PH, MB) with HMS-4 criteria.[1] Developmental assessments included a review of prospective developmental records from parents, health visitors, and general practitioners. Four children did not undergo psychiatric assessment in hospital; all had been assessed professionally elsewhere, so these assessments were used as the basis for their behavioural diagnosis.

After bowel preparation, ileocolonoscopy was performed by SHM or MAT under sedation with midazolam and pethidine. Paired frozen and formalin-fixed mucosal biopsy samples were taken from the terminal ileum; ascending, transverse, descending, and sigmoid colons, and from the rectum. The

Figure 1: Published journal article that linked measles vaccines to autism.

Statistics/Biostatistics

Being statistically literate allows you to...

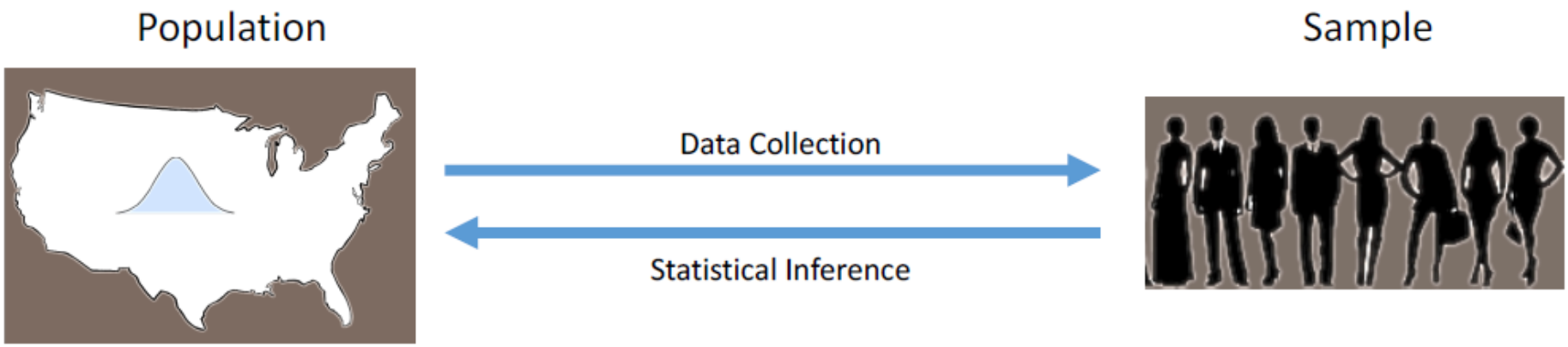
- Sift through the seeming randomness of the world to find potentially meaningful patterns.
- Critically interpret scientific research papers and news articles
  - What is reliable? What is not?
- Structure effective scientific research
- Interpret data and present in a way that non-statisticians can understand
- Collaborate with a broad range of scientific investigators

The world is your oyster!

Populations and Samples

Populations versus Samples

- **Population**: every member of the group of interest
  - Ex: women in America
- In a perfect world, we would collect data on every subject in a population
- This is rarely possible. Instead, we draw a **sample** from the population.
  - Ex: randomly survey 20 women from every state in America.
- We apply **statistical inference** to the sample to make conclusions (with a certain level of "confidence") about the population.



Parameters versus Statistics

- **Parameters**: Summarize characteristics of an entire population
  - Fixed, usually unknown numerical value
  - Ex: population mean (μ), population variance (σ^2)

- **Statistics**: Summarize charactersitics of a sample
  - Changes with sample, known number
  - Ex: sample mean ($\bar{x}$), sample proportion ($\hat{p}$), sample standard deviation (s)

Descriptive versus Inferential Statistics

- **Descriptive Statistics:** Describe the sample
  - Do not allow us to make conclusions about population
  - Often, use measures of central tendency (sample mean, sample median) and measures of spread (sample standard deviation)

- **Inferential Statistics:** Use sample to make conclusions about population
  - Ex: estimation of population parameters (with some "confidence"), statistical hypothesis testing

Point versus Interval Estimates

In inferential statistics, we use the sample to estimate population parameters with both a single value and a range of possible values:

- **Point estimate**: a single value of a statistic.
  - Ex: $\bar{x}$ is a point estimate of μ
- **Interval estimate**: Two numbers, between which a population parameter is likely to be between.
  - Usually have some level of "confidence" (we will discuss confidence intervals later)
  - Ex: (a, b), where $a < \bar{x} < b$, is an interval estimate for μ. μ is likely to be in this interval.