

# Bioinformatics and the investigation of cholera

George Bell, Ph.D.

Bioinformatics and Research Computing

# Bioinformatics

- Using computers to organize and analyze biological information
- Common types of biological information
  - DNA, protein, and genomic sequences
  - Large-scale quantitative data: microarrays, protein-protein interactions
- Sometimes called “computational biology”

# Major areas of BaRC work

- Provide lots of software (desktop, web, and Unix) and databases to all scientists
- Train people to use this software
- Teach classes in bioinformatics theory and practice
- Consult and collaborate with scientists on their specific research interests
- Create custom software and databases

# Learning bioinformatics

- Bioinformatics gets people from all directions and backgrounds
- Most helpful academic subjects:
  - Biology
    - Esp. molecular biology, genetics, biochemistry
  - Computer science, mathematics, and/or statistics
- Experience in a lab that does what you might like to do

# Cholera

- 1853: What is cholera? ...all is darkness and confusion, vague theory, vain speculation...

-- The Lancet

- “The symptoms are primarily seated in the alimentary canal, and all the after-symptoms of a general kind are the results of flux from the canal”

-- Richardson, “Snow on Cholera”

- Death in a few hours to a week

William Farr  
1807-1883



Well-known  
epidemiologist

# London 1848-9 cholera epidemic: 1.9 million deaths

- Deaths per 10,000 people
- Elevation above the Thames
- Persons per acre
- Persons per house
- Average annual value of house
- Average annual value of house per person
- Poor rate
- Water supply

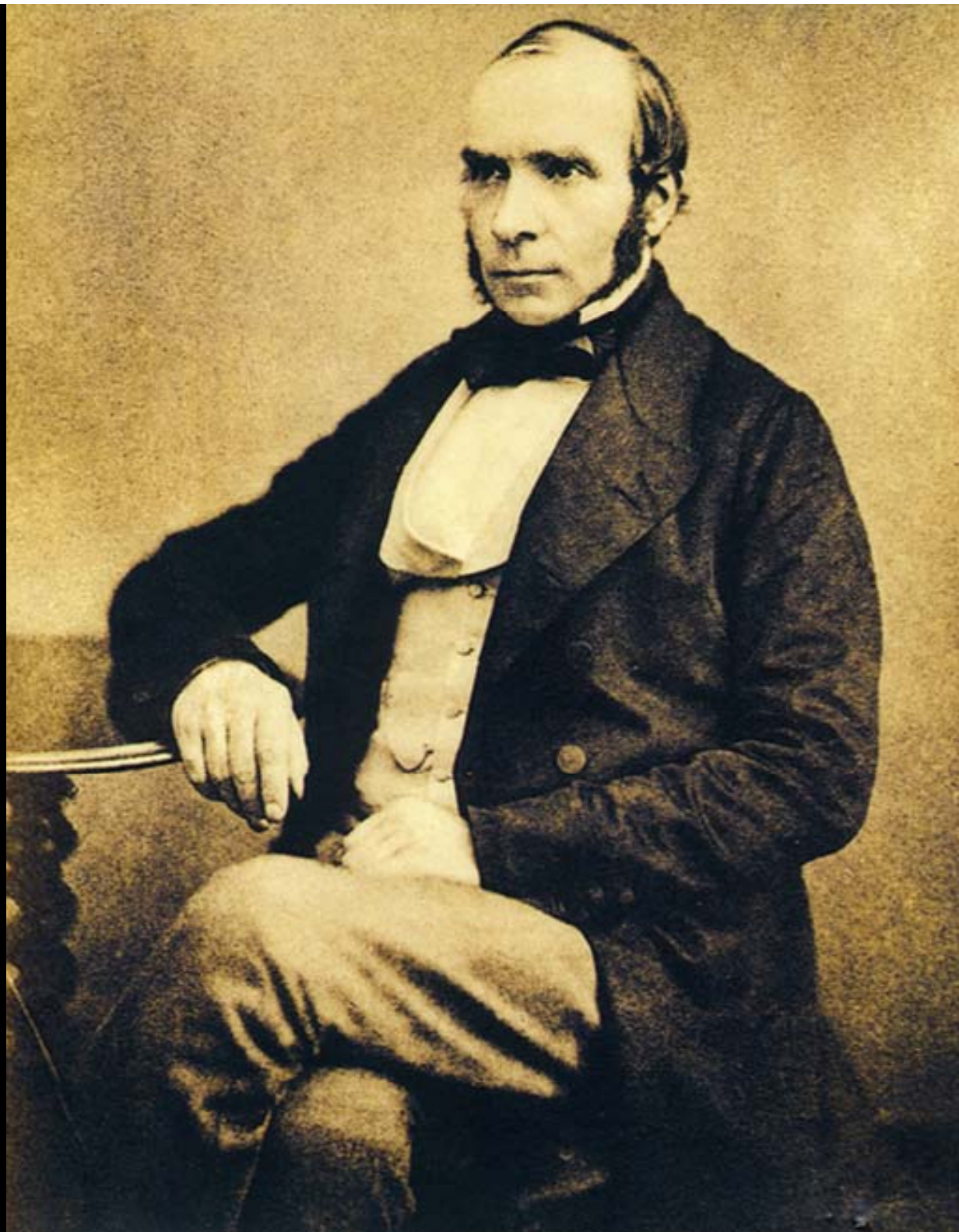
# Summary of Farr's analysis

The elevation of the soil in London has a more constant relation with mortality than any other known element

- Note: He was wrong



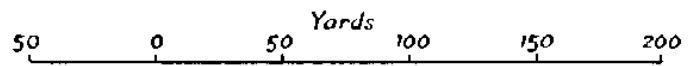
John Snow  
1813-1858



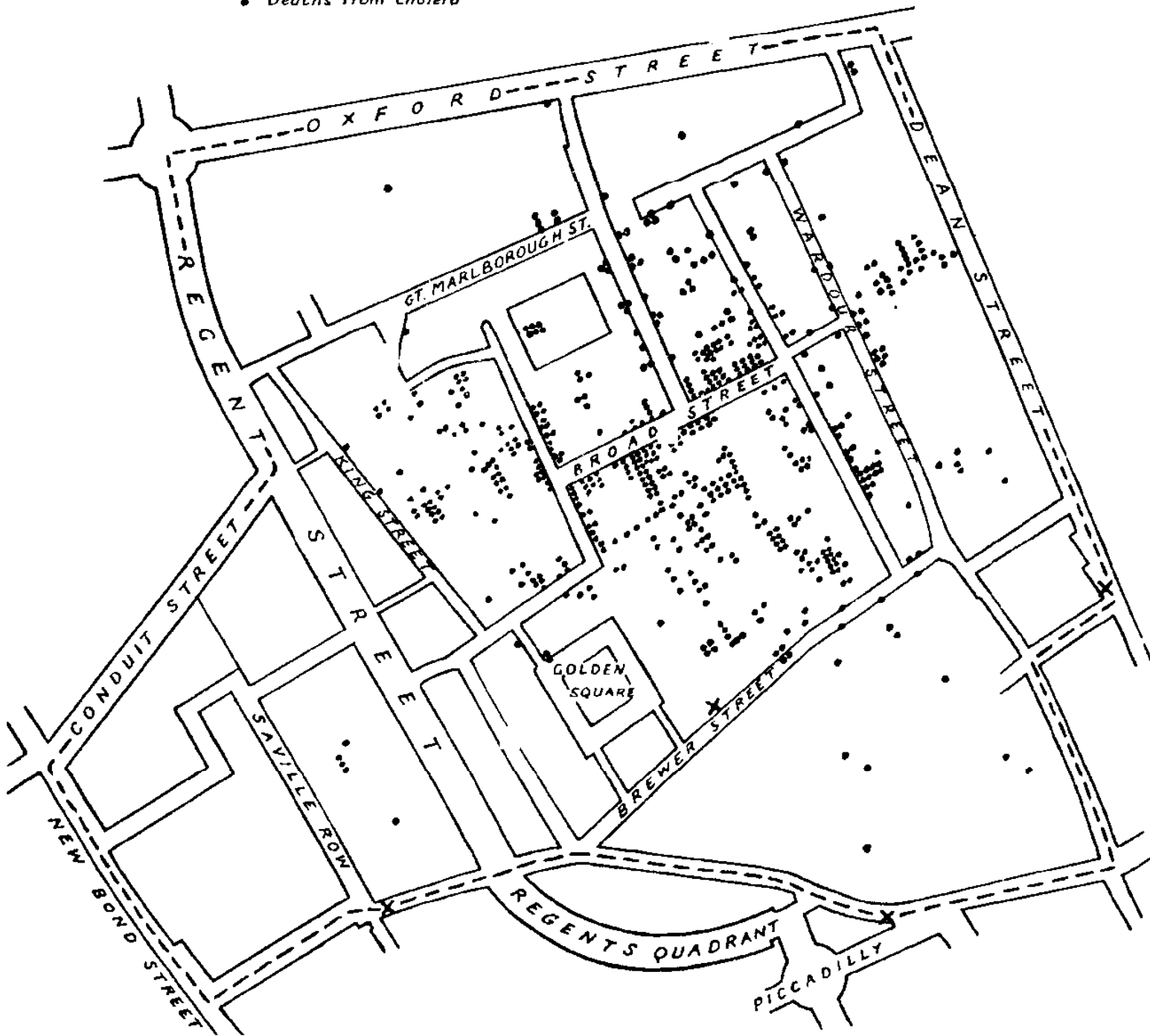
Up-and-coming  
doctor

# London 1854 cholera epidemic

- Since this was an intestinal disease, Snow already felt confident it was due to something that was eaten or drunk.
- While the epidemic was in progress, he mapped all deaths to London locations.

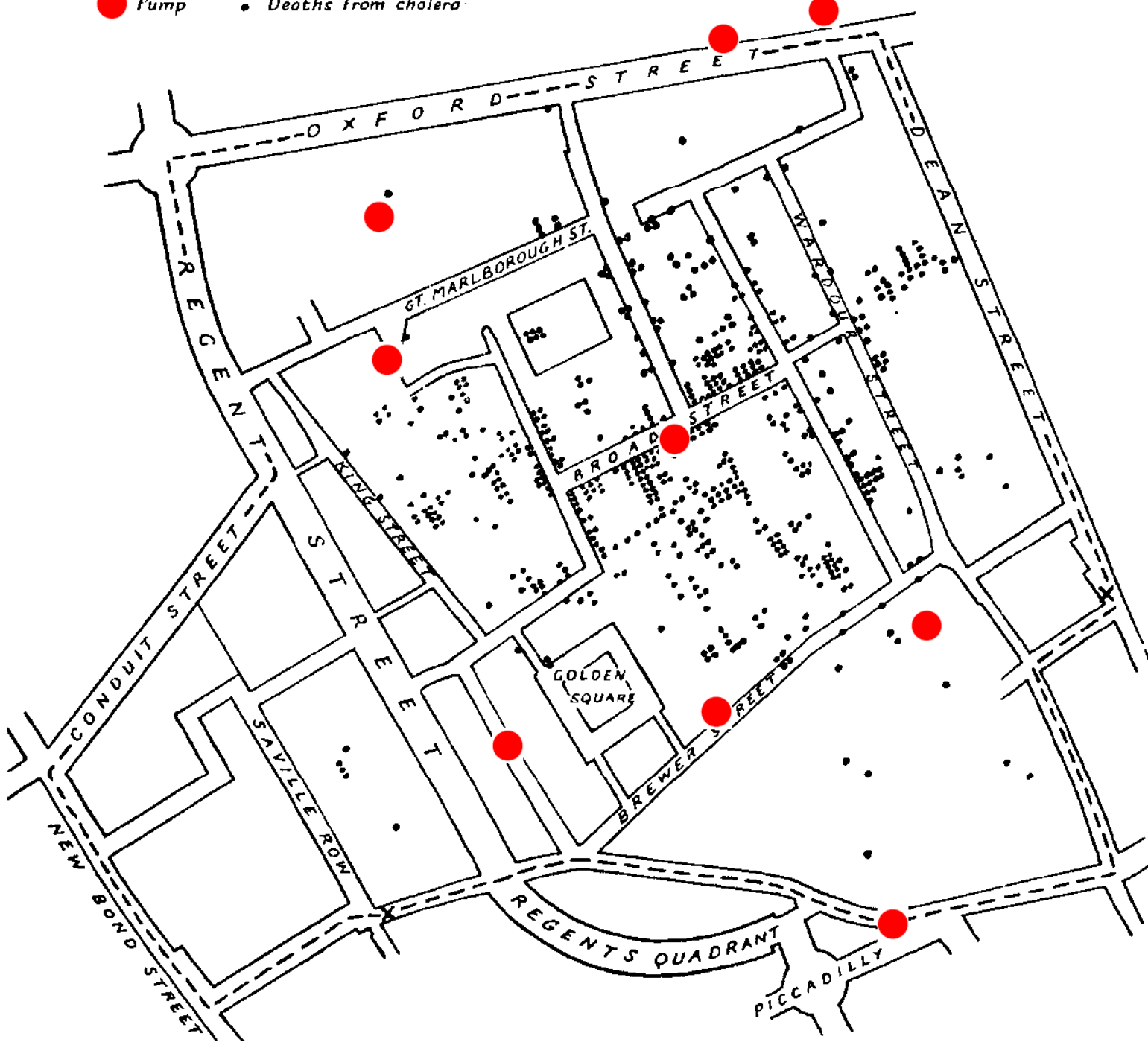


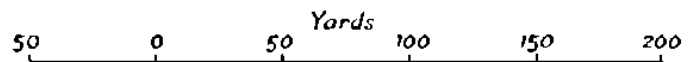
• Deaths from cholera



50 0 50 100 150 200  
Yards

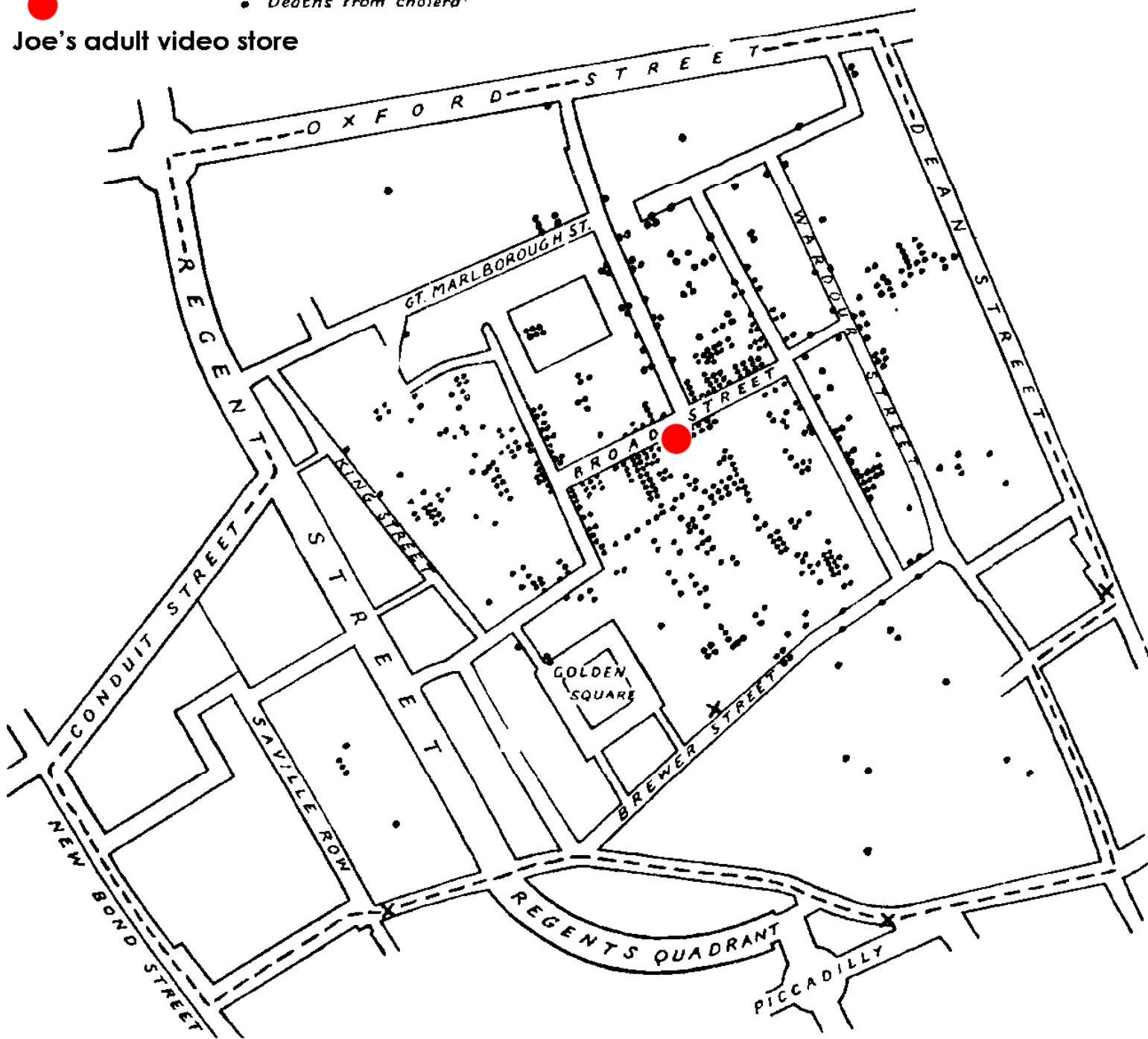
● Pump    • Deaths from cholera





• Deaths from cholera

Joe's adult video store

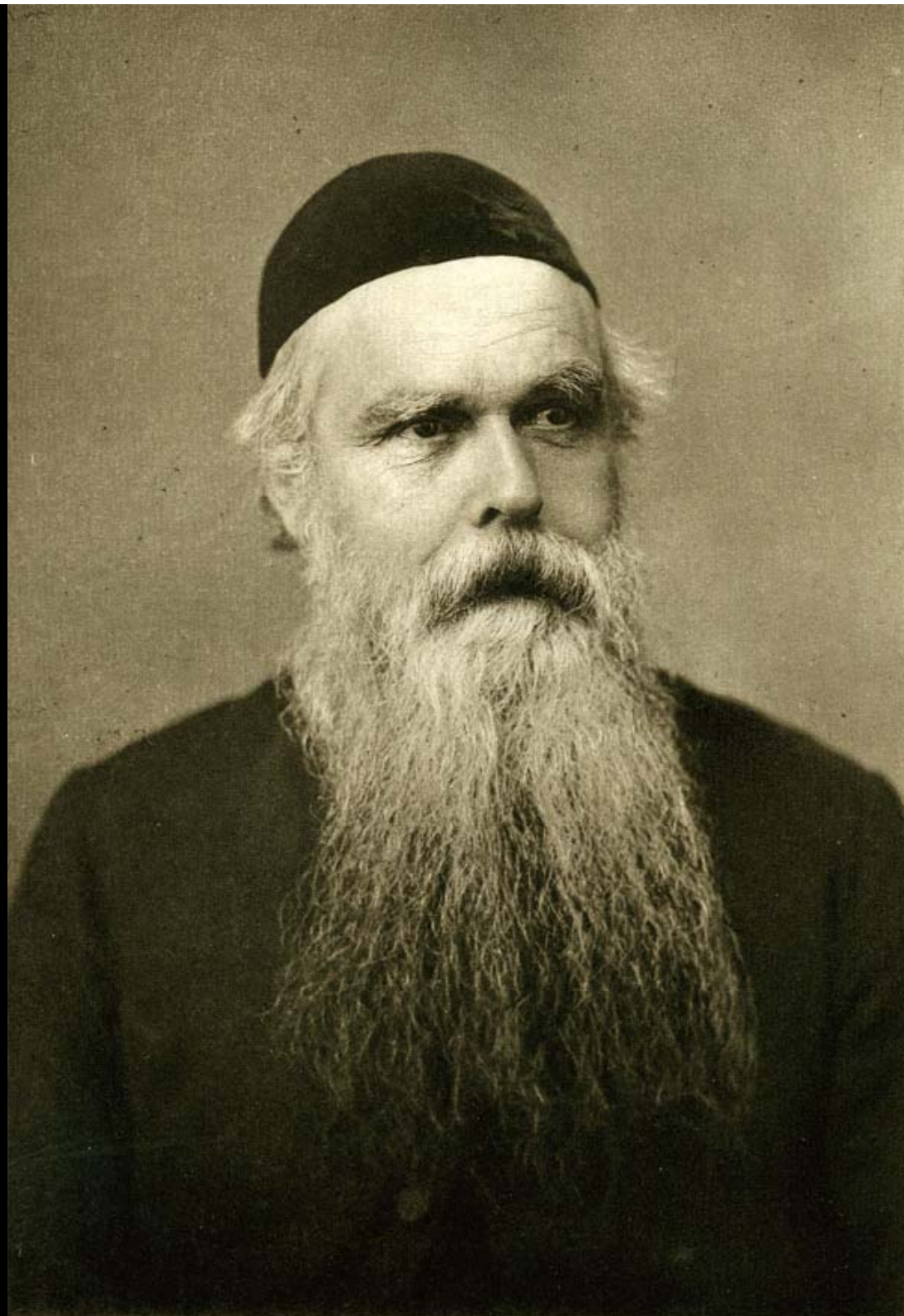


# The Broad Street pump

- Hypothesis: The pump was the origin of the cholera outbreak
- People listened to Snow but weren't convinced.
- What experiment could be designed to test this hypothesis?
- Note: He was right



Henry Whitehead  
1825-1896



minister

# Whitehead's interviews

- Set out to disprove Snow's ideas
- The source of the outbreak was a mother washing an infected baby's diapers on August 24, 1854
- Also explained unexpected observations
- Came to accept Snow's ideas



Robert Koch  
1843 - 1910



Official  
discoverer  
of  
*Vibrio  
cholerae*

# Benefits of modern biology

- Genome sequences
- Protein sequences
- Many protein structures
- Physiology of cholera infections

# Escherichia coli



Department of Health and Human Services  
Centers for Disease Control and Prevention

[CDC en Español](#)

Search:

## *E. coli* Outbreak Investigations

**Note:** the updates below have not been revised since their original release, and the content they contain may no longer be up to date.

### *Escherichia coli* O157:h7 (Topp's Ground Beef Patties)

- ◆ [October 26, 2007](#)

### *Escherichia coli* O157:h7 (Taco Bell)

- ◆ [December 14, 2006](#)
- ◆ [December 13, 2006](#)
- ◆ [December 12, 2006](#)
- ◆ [December 11, 2006](#)
- ◆ [December 10, 2006](#)
- ◆ [December 9, 2006](#)
- ◆ [December 8, 2006](#)
- ◆ [December 7, 2006](#)
- ◆ [December 6, 2006](#)

### *Escherichia coli* O157:h7 (Fresh Spinach)

- ◆ [October 6, 2006](#)

# Protein structure demo

- Goal: Analyze the structure and function of the cholera and E. coli toxins