Existing Literature

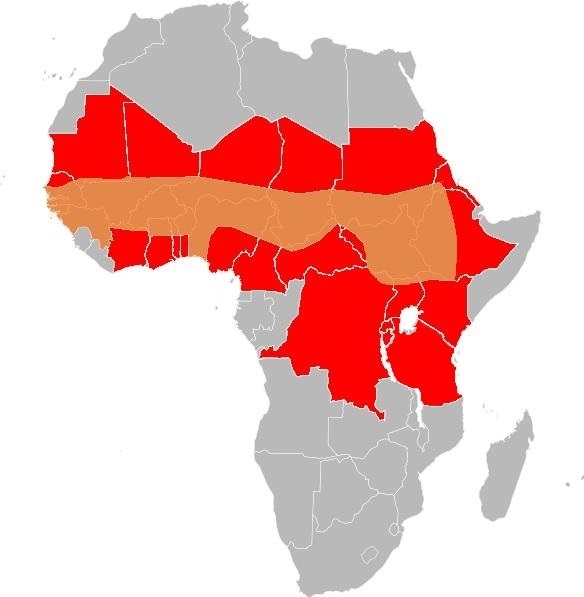
* <http://www.who.int/bulletin/volumes/90/6/11-086009.pdf>
* **Research paper that models meningitis outbreaks using a Bayesian network approach**
* <https://www.hindawi.com/journals/cmmm/2018/2657461/>
* <https://www.wmo.int/pages/prog/arep/wwrp/new/documents/3_9_2_MERIT_Strategic_Review.pdf>
* Report that analyzes techniques and strategies for monitoring and controlling meningitis outbreaks
* <https://www.researchgate.net/publication/263741627_Climate_Change_and_Cerebrospinal_Meningitis_in_the_Ghanaian_Meningitis_Belt>
* <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6292165/>
* <http://www.ajtmh.org/content/journals/10.4269/ajtmh.2012.11-0597#html_fulltext>
  + Model uses social and news media enable estimation of epidemiological patterns of Cholera (what inspired my initial research question)

Resources

* <http://www.menafrinet.org/en-us/Resources>
* <https://www.who.int/emergencies/diseases/meningitis/epidemiological/en/>
  + Weekly surveillance reports by the WHO
* Excailbur (tool for extracting tables from PDF) and Ghostscript
* <https://hackernoon.com/an-open-source-science-tool-to-extract-tables-from-pdfs-into-excels-3ed3cc7f22e1>
* <https://camelot-py.readthedocs.io/en/master/user/advanced.html>

Data? <http://iridl.ldeo.columbia.edu/SOURCES/.LSTM/.MFAP/.molesworth_etal_2003/>

The Meningitis Belt



Composed of 26 countries in sub-Saharan Africa

* Very fast / hard to control out-breaks – can kill up to half the people who show symptoms

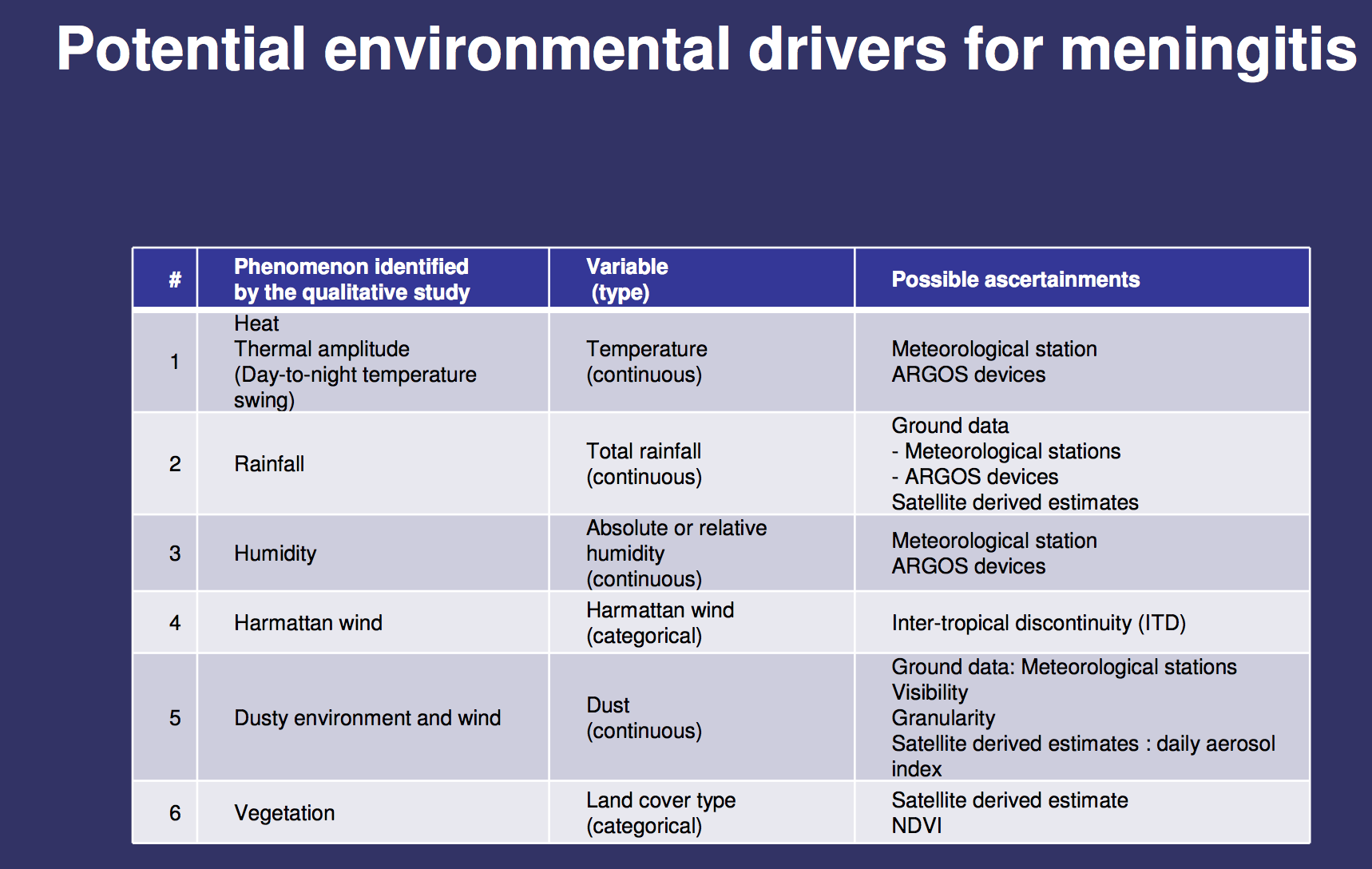
**Q. What is meningitis?**

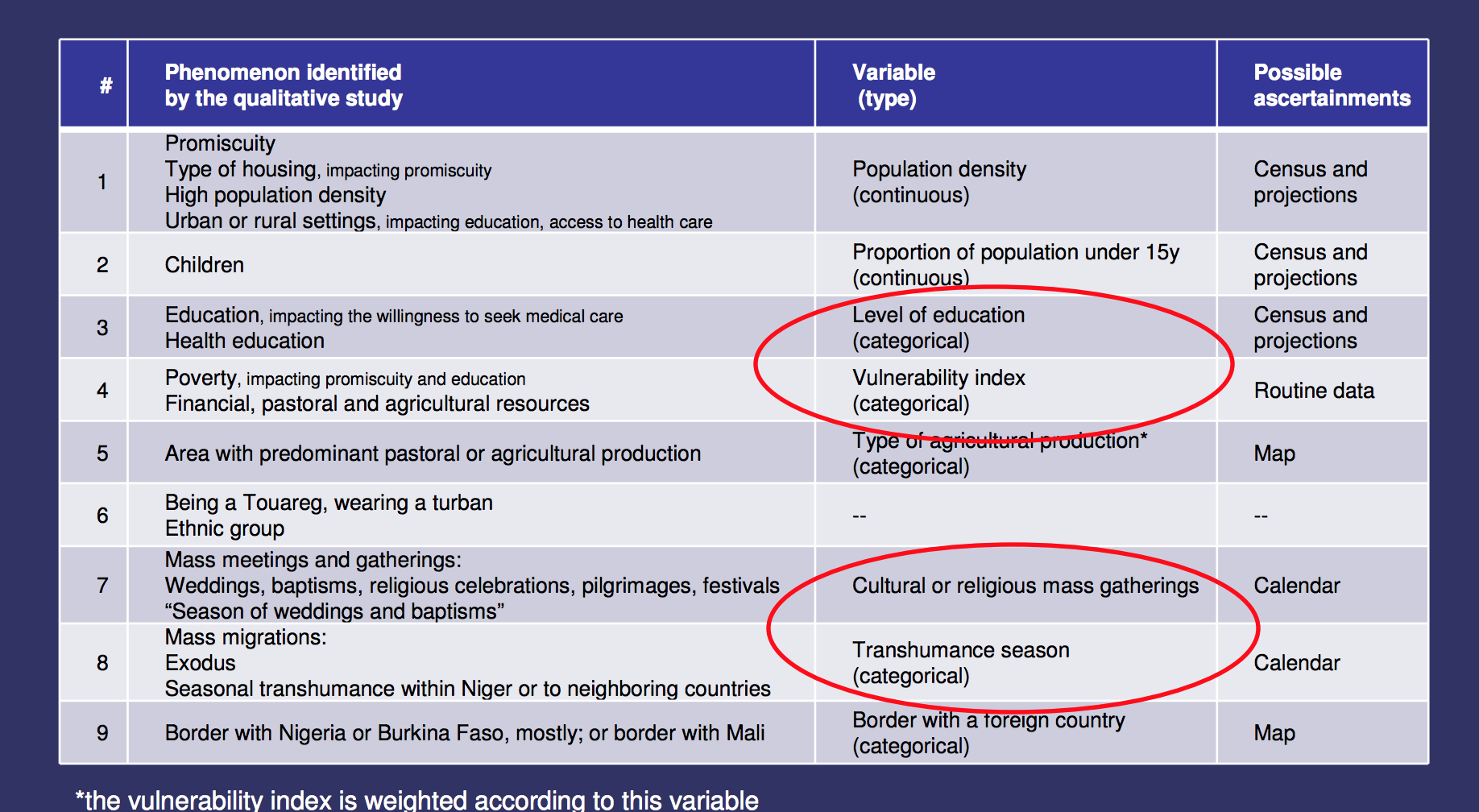
* Inflammation of membranes near brain and spinal cord
* Meningitis is found in two forms: bacterial and viral – will focus on bacterial which is significantly more dangerous
* Within 2 days, if not treated, can lead to severe brain damage and in some instances, death

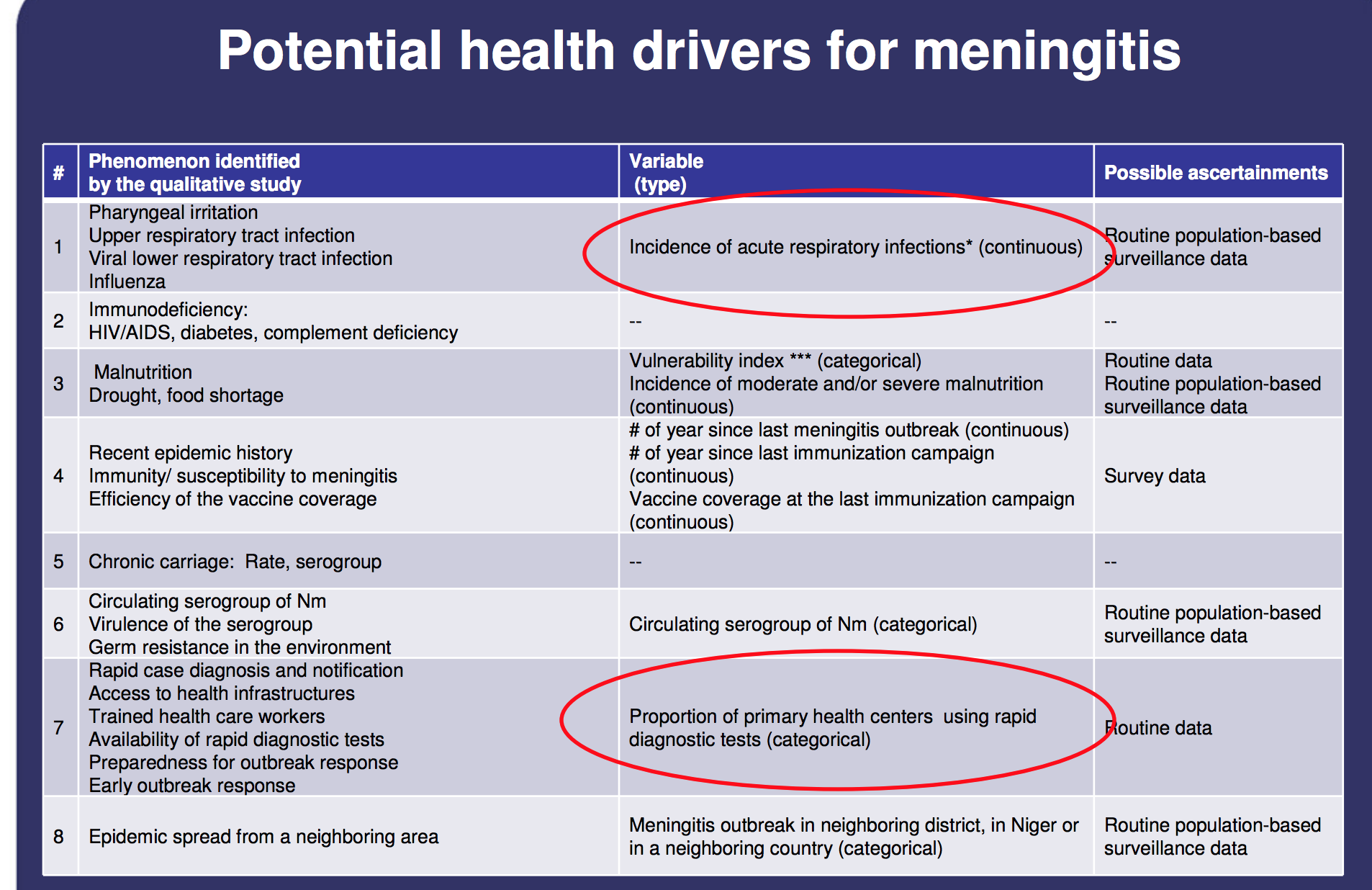
**Initially Identified Potential Risk Factors**

* Immunological susceptibility
* Special climatic conditions (dry season/drought/dust storm)
* Population density
  + More people in a crowded space (military camps, pilgrimages…) help spread the disaease
* Socioeconomic status
  + Poor living conditions are correlated with higher incidence of the disease
* Respiratory tract
  + Can aid spread of an outbreak
* Age (decreases with age)

**More Comprehensive List:**

****





**Algorithm Side**

* Bayesian networks
* Dynamic Bayesian Networks
* Markov chains and hidden Markov models

**Epidemiology Side**

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3196833/pdf/ukmss-36404.pdf>

To monitor and survey disease, two broad methods

1. Statistical methods
   1. Modelling based on statistical relationships established between past case and environment
2. Biological methods
   1. Creating equations which encapsulate the transmission
   2. Require details on factors considered to be important in transmission process
      1. Which is why we often use statistical models since we don’t know all of the biological factors that impact the transmission
   3. <https://www.hindawi.com/journals/cmmm/2018/2657461/>

* Important to combine both statistical and biological approach

\*\* any prediction of risk should include some kind of estimate of its reliability – if critical ensures resources are quickly mobilized

Types of Early Warning Systems

1. Reportorial System
   1. Uses data that gives us number of laboratory/hospital based cases of a certain disease
2. Risk Mapping System
   1. Both space and time series used
   2. \*Assume that patterns from the fast can be projected to the future to make predictions about what occurs
   3. Look into the **DEWS forecasting model**
      1. Applied temporal Fourier analysis of de-trended time series splits to model within-year and across year variations
3. Environmental Early Warning Systems

Q. What makes a good prediction?

* Accuracy
* Geographical scale
* Temporal duration
* Should be linked with response initiatives so that they can update based on these actions

**Bayes Theorem**

<https://machinelearningmastery.com/what-is-a-hypothesis-in-machine-learning/>

<https://machinelearningmastery.com/bayes-theorem-for-machine-learning/>

Quantifies the actual accuracy of tests by correcting for the skew of false positives

Intuitively – p(hypothesis | evidence) = p(true positive) / p(positive)

* P(true positive) = p(evidence | hypothesis) \* p(hypothesis)

Explanations words – the plausibility for a belief indicates the degree to which your belief is explained only by the evidence and not by anything else (i.e. false positives)