**NURS 8310 Week 2 Assignment: Strengths and Limitations of Secondary Data Sources**

Student Full Name

Institution Affiliation

Course Full Title

Instructor Full Name

Due Date

**NURS 8310 Week 2 Assignment: Strengths and Limitations of Secondary Data Sources**

Nursing researchers rely massively upon primary and secondary data sources to answer foreground questions, support research hypotheses and inform evidence-based practices. Secondary sources exist in many forms, including quantitative meta-analysis articles, newsletters, government publications, practice guidelines, and standards, systematic review articles, and book reviews. These sources focus on primary data and provide in-depth analysis, commentary, or discussion regarding original research works. Secondly, secondary evidence sources summarize, compare, and evaluate primary information to draw conclusions and establish the extent and the current state of knowledge in a discipline or topic of scholarly exploration. Wickham (2019) contends that secondary data investigators must familiarize themselves with their research areas and identify datasets that provide valid and relevant information on the research area.

Information validity and reliability are profound considerations when selecting secondary evidence sources. For instance, nursing researchers can access reliable and valid secondary information from reputable databases, including the Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO). These free-to-access databases contain datasets from multiple healthcare research areas. Consequently, the focus of this paper is to identify the selected health problem captured in these databases, datasets regarding the problem, and variables of interest, and assess the validity of each data set. Further, the paper elaborates on the potential challenges encountered in identifying a proper dataset or securing permission to use it.

**Brief Identification of a Population Health Problem**

Tobacco use poses a significant population health challenge, considering its association with multiple life-threatening conditions. According to the Centers for Disease Control and Prevention [CDC] (2021), tobacco use through cigarette smoking is a global healthcare problem that accounts for higher mortality rates and various comorbidities. For instance, more than 480000 Americans succumb to smoking-related diseases annually. Such statistics render tobacco use as the leading cause of death above HIV/AIDS, alcoholism, motor vehicle injuries, and firearm-related incidents (Centers for Disease Control and Prevention, 2021). Regarding tobacco use as a risk factor for other life-threatening conditions, cigarette smoking alone accounts for approximately 90% of all lung cancer deaths and 80% of all mortalities associated with chronic pulmonary obstructive disease (COPD).

Besides exacerbating the rates of premature deaths, tobacco use, especially cigarette smoking exposes people to increased health risks. For example, cigarette smokers are more likely to develop heart disease, stroke, and lung cancer than nonsmokers. Other health issues associated with tobacco use are preterm delivery, stillbirth, low birth weight, ectopic pregnancy, poor dental health, a high risk of cataracts, type 2 diabetes mellitus (T2DM), rheumatoid arthritis, and the overall decrease in immune function (Centers for Disease Control and Prevention, 2021). Based on alarming statistics on tobacco use prevalence, mortality rates, and comorbidities, nursing researchers perceive this public health problem as a sensitive research topic. As a result, research works focus on identifying the variables and determinants of tobacco use, the plausibility of tobacco use cessation programs, and the effectiveness of regulatory policies in reducing the risk and rates of tobacco use among susceptible groups.

**The Selected Dataset**

The selected datasets are consistent with the identified public health problem (tobacco use) because they provide information and statistics on tobacco use prevalence and trends over time. The first selected dataset is "BRFSS Prevalence and Trends Data: Tobacco Use-Four Level Smoking Data for 1995-2010." This dataset is available on the Centers for Disease Control and Prevention website and provides data on weighted percentages of people based on various population characteristics, including people who smoked every day, some days, former smokers, and those who never smoked cigarette across the stated time (1995-2010) (Centers for Disease Control and Prevention, 2013). To aid visualization, the dataset comprises a tabular preview consisting of columns for the year, state, and stipulated population characteristics (regular smokers, occasional smokers, former smokers, and non-smokers).

The second dataset identified in this assessment is the World Health Organization's noncommunicable diseases data portal which provides statistics on various chronic conditions, including cancer, cardiovascular diseases (CVDs), chronic respiratory diseases (CRDs), diabetes, alcoholism, and tobacco use. After navigating the World Health Organization's portal for non-communicable diseases, I selected tobacco use as the priority condition. Consequently, I accessed a simple dataset, "Global-Current tobacco use, adults aged 15+ (2019)." This dataset provides gender-based statistics for tobacco use rates globally and regionally (World Health Organization, n.d.). Regions captured in the dataset are regions of the Americans, South-East Asian Region, European Region, African territories, Eastern Mediterranean Region, and Western Pacific Region. It is possible to Identify the rates of tobacco use in these regions by studying the map and the subsequent colors that represent prevalence rates.

**Identification of Variables in Each Dataset**

The identified datasets provide information regarding trends in tobacco use and the rates of tobacco use in American states and other regions. As a researcher, I can use these datasets as ideal secondary data sources when researching the trends and prevalence of tobacco use. Although the two datasets are ideal secondary evidence sources, they do not identify variables that explain the disproportionate prevalence of tobacco use among adults in different regions. According to Gutema et al. (2021), tobacco use is a multifactorial healthcare problem whose prevalence depends on various determinants. For example, the independent variables in tobacco use include the target population's socio-demographic status, and demographic aspects like sex, age, residency, marital status, educational level, wealth status, and occupational status. Also, people's mental health status and the prevalence of psychological problems like stress can account for the high prevalence of tobacco use. Therefore, it is essential to study the association between these independent variables and tobacco use as a dependent variable.

**Assessing the Validity of Each Dataset**

Data validity is a profound consideration when appraising evidence sources because it explains how well the documented findings represent true findings among similar individuals outside the study (Patino & Ferreira, 2018). In primary studies, data validity exists in internal and external domains. Internal validity entails the extent to which the results represent the truth in the studied population and the absence of methodological errors. Conversely, external validity explains the plausibility of applying the results to similar patients in different settings (Patino & Ferreira, 2018). Data validity may entail the meaningful and appropriate interpretation of data obtained from primary sources in secondary data sources.

The selected datasets pass various criteria for data validity. Firstly, they are published in reputable and reliable databases. The Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO) provide information and justifiable statistics on multiple health topics of national and global interest. These agencies have comprehensive systems and frameworks for collecting primary data and developing accessible datasets for individual and organizational use. Before publishing information in these databases, data scientists use all quality, validity, and reliability measures to identify and correct methodological errors. Secondly, these databases are ideal sources for secondary studies. Healthcare researchers partner with CDC and WHO to publish their scholarly works, including journal articles. These empirical studies undergo the peer-reviewing process to confirm if they meet all standards of scholarly publications. Based on these comprehensive data collection and publication processes, it is valid to argue that the identified datasets meet all data validity criteria.

**Potential Challenges in Identifying a Proper Dataset or Securing Permission**

Many secondary evidence sources provide structured health data that is publicly available and free to access. According to Areco et al. (2021), the Centers for Disease Control and Prevention (CDC) is among the reputable databases that provide secondary health data for research and organizational use. However, these databases can provide restricted access data that require authorization and subscription. Data authorization and subscription are among the problems encountered by researchers when identifying proper datasets for research topics and questions. Secondly, researchers can face challenges in identifying ideal datasets for research due to a lack of standardization in the data format, discontinuity in data collection over time, variation in coverage, and location-based data limitations (Areco et al., 2021). Thirdly, some databases require researchers to contact data owners or seek permission before accessing datasets. These issues are problematic because researchers are vulnerable to the delayed response by data owners and the potential restrictions due to data privacy protocols. Researchers should be aware of these constraints and plan effectively before embarking on research endeavors.

**Conclusion**

Secondary evidence sources provide commentary, discussion, and analysis of primary sources to draw conclusions and present the current state of empirical knowledge in a topic or discipline. Examples of secondary data sources include book reviews, survey articles, government publications, quantitative meta-analysis articles, entries in medical encyclopedias, and practice guidelines. The selected public health problem for this assessment is tobacco use. Therefore, the identified datasets from the Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO) provide trends and prevalence of tobacco use in the United States and globally. Although these datasets are valid and accessible, many databases provide restricted access data, require researchers to contact data owners and incorporate rigorous authorization processes according to data privacy and security protocols. Researchers face these challenges when accessing secondary data sources. Consequently, they should be aware of these drawbacks and plan effectively before commencing research activities.

**References**

Areco, K. N., Konstantyner, T., Bandiera-Paiva, P., Balda, R. C. X., Costa-Nobre, D. T., Sanudo, A., Kiffer, C. R. V., Kawakami, M. D., Miyoshi, M. H., Marinonio, A. S. S., Freitas, R. M. V., Morais, L. C. C., Teixeira, M. L. P., Waldvogel, B., Almeida, M. F. B., & Guinsburg, R. (2021). Operational challenges in the use of structured secondary data for health research. *Frontiers in Public Health*, *9*, 642163. <https://doi.org/10.3389/fpubh.2021.642163>

Centers for Disease Control and Prevention. (2013). BRFSS prevalence and trends data: Tobacco use – Four-level smoking data for 1995-2010. https://data.cdc.gov/Smoking-Tobacco-Use/BRFSS-Prevalence-and-Trends-Data-Tobacco-Use-Four-/8zak-ewtm

Centers for Disease Control and Prevention. (2021). Health effects of cigarette smoking. *Centers for Disease Control and Prevention.* [https://www.cdc.gov/tobacco/data\_statistics/fact\_sheets/health\_effects/effects\_cig\_smoking/index.htm#](https://www.cdc.gov/tobacco/data_statistics/fact_sheets/health_effects/effects_cig_smoking/index.htm)

Gutema, B. T., Chuka, A., Ayele, G., Estifaons, W., Melketsedik, Z. A., Tariku, E. Z., Zerdo, Z., Baharu, A., & Megersa, N. D. (2021). Tobacco use and associated factors among adults reside in Arba Minch health and demographic surveillance site, southern Ethiopia: A cross-sectional study. *BMC Public Health*, *21*(1). <https://doi.org/10.1186/s12889-021-10479-4>

Patino, C. M., & Ferreira, J. C. (2018). Internal and external validity: Can you apply research study results to your patients? *Jornal Brasileiro de Pneumologia*, *44*(3), 183–183. <https://doi.org/10.1590/s1806-37562018000000164>

Wickham, R. (2019). Secondary analysis research. *Journal of the Advanced Practitioner in Oncology*, *10*(4), 395–400. <https://doi.org/10.6004/jadpro.2019.10.4.7>

World Health Organization. (n.d.). Non-communicable diseases data portal. Retrieved March 13, 2023, from <https://ncdportal.org/>