**Student Handout A: Calculating Lifetime Cancer Risk Resulting from DNA Replication**

*Learning objectives:*

* Utilize experimental design to propose a scientific question relating one variable to another, generate a hypothesis, and formulate a method to test the hypothesis.
* Express numbers in scientific notation and percentage.
* Evaluate a Pearson correlation coefficient (r) to identify relationships between variables.
* Illustrate correlation between variables by generating scatterplots.
* Explain the role of stem cells in the development and maintenance of tissues.
* Assess the effect mutations in the replication machinery would have on DNA replication.
* Relate how errors in DNA replication could affect lifetime cancer risk.

Cancer is a disease in which some of the body’s cells begin to divide without stopping and spread into surrounding tissues.Cancer risk has been demonstrated to be associated with:

* + inherited genetic variation/mutations (5 to 10% of cancers have a heritable component)
	+ environmental risk factors (i.e. smoking, alcohol use, ultraviolet light, human papilloma virus (HPV))

These factors alone, however, fail to explain why some of the body’s tissue types give rise to human cancer millions of times more often than others- even in cases where hereditary and environmental risks are equal. In this paper, researchers investigated how a third factor—chance genetic mutations that occur during stem cell division—may contribute to the overall rate of cancer incidence in each tissue. To understand this, we must first understand more about stem cells, and what must happen as they prepare to divide.

*Stem Cells:*

* Stem cells often comprise only a small fraction of the total cells in a given tissue.
* Stem cells are responsible for the development and maintenance of tissues, and are capable of dividing via mitosis over a long period of time.

*DNA Replication:*

* DNA replication is required for all cells to divide, occurring prior to both mitosis and meiosis.
* DNA replication results in an error rate of one in 105 nucleotides.

*Reducing Errors Generated in Replication*

* DNA proofreading reduces this error rate (mutation rate) to only one in 1010 nucleotides. This mutation rate seems to be identical for all cell types.

DNA Proofreading: Mismatch Repair:



Use this resource to learn more about DNA proofreading and repair: https://www.khanacademy.org/science/biology/dna-as-the-genetic-material/dna-replication/a/dna-proofreading-and-repair

1. In your own words, generate a scientific question that relates DNA replication to cancer risk.
2. Propose a hypothesis to answer your question. [What will you measure (that is, what’s the dependent variable)? What will you manipulate (independent variable)?] Write your statement of a testable hypothesis here. The hypothesis will state a predicted relationship between the independent variable and the dependent variable.
3. What are the variables being analyzed in your analysis? What are your controls?
4. Suggest one way to test your hypothesis.

*In 2015, C. Tomasetti & B. Vogelstein published a study asking if the overall risk of cancer over a person’s lifetime for a given tissue increase for tissues with a greater number of stem cell divisions (and therefore greater amount of DNA replication). The data in figure 1 is from this paper. Use Figure 1 to answer questions 8-9.*

**

*Figure 1*

1. What trends do you notice in this figure?
2. What do the trends in this figure tell you about the relationship between stem cell divisions and cancer risk?
3. Given:
* The lifetime risk of all pancreatic cancers is 1.49%.
* Endocrine (islet cell) carcinomas account for about 1.3% of all pancreatic cancers.
* The multiplication rule is used to assess risk when one event is dependent on the occurrence of another.
* The addition rule is used to assess the risk of any one of two or more events occurring.
1. Calculate the lifetime risk of pancreatic islet cell carcinoma.
2. Which rule is used for this calculation- multiplication or addition?
3. Express lifetime risk in scientific notation.
4. Express lifetime risk as a percentage.

8. Correlation measures the strength and the direction of the linear relationship between two variables. This can be measured by a Pearson correlation coefficient (r). The greater the absolute value of r, the stronger the linear relationship between the two variables.

* + When r = +1, the variables are perfectly linearly correlated in the same direction. When one variable increases, the other increases.
	+ When r = -1, the variables are perfectly linearly correlated in the opposite direction. When one variable increases, the other decreases.
	+ When r = 0, no correlation between the two variables.

*Use the space given to generate the following scatterplots:*

1. Generate a scatterplot with at least 10 data points given r = +1.
2. Generate a scatterplot with at least 10 data points given r = -1.
3. Generate a scatterplot with at least 10 data points given r = 0.

a. b. c.

9. Generate a statement that interprets the meaning of each of the following r-values. In your statement, indicate both the **strength** of the correlation as well as the **direction** of the correlation.

1. r = -0.2
2. r = +0.9
3. r = +0.4

10. What does the phrase “positively correlated” tell us?

11. Write a statement that reflects what you can conclude about the relationship between the number of stem cell divisions a tissue type incurs and the lifetime cancer risk?

12. Why do you think that there is a positive correlation between these two variables?

*Reflection*

13. State one thing you learned about experimental design in this exercise.

14. State one thing you learned about DNA replication in this exercise.

15. Write a specific statement of one way measuring Pearson's correlation coefficient values can be used.

16. For you, what was the most difficult thing about this exercise?

**References**

1. Tomasetti, C. & Vogelstein, B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* **347**, 78–81 (2015); <http://science.sciencemag.org/content/347/6217/78.full>
2. OpenStax CNX; https://cnx.org/contents/esgfrPlv@3/Accessory-Organs-in-Digestion-
3. HHMI Biointeractive; <https://www.hhmi.org/biointeractive/data-points>
4. Khan Academy; https://www.khanacademy.org
5. Math Bench Biology Modules; <http://mathbench.umd.edu/modules/misc_scaling/page07.htm>