

# Zika virus-related Photo sharing on Pinterest and Instagram

## Appendix B

### Methods

#### Process of coder training

We took the following steps in our training and coding procedure:

1. We provided the two coders with Training session #1.
2. The coders coded the first Pinterest training set (721 Pinterest photos) independently.
3. We provided the two coders with Training session #2.
4. The coders then coded the second Pinterest training set (741 Pinterest photos) and the Instagram training set (741 Instagram photos) independently.
5. We provided the two coders with Training session #3.
6. The two coders coded the final (research) data sets independently.
7. The co-first author resolved all discrepancies between the coders: this final coded sample was used in our statistical analysis.

#### Data collection for training data sets

*Pinterest training sets.* Pinterest training data sets (with keywords: “zika” AND “virus”) were captured using a browser add-on called “Full Page Screen Capture.”<sup>1</sup>

- The first Pinterest training set contained 721 Pinterest photos that were retrieved on January 29, 2016.
- The second Pinterest training set contained 741 Pinterest photos that were retrieved on February 8, 2016.

*Instagram training set.* A total of 5043 Instagram photos (with the hashtag “#zikaavirus”) were captured via Instagram Application Programming Interface on February 8, 2016. Of these 5043 photos, a random sample of 741 was selected as the Instagram training set.

#### Explanatory notes about Instagram Application Programming Interface

KCC wrote the Python code. NS used the code to retrieve the Instagram photos to a server at Athens, Georgia, USA. We captured Instagram photos only, without their tag/caption text.

#### Explanatory notes about embedded words

For embedded words, we referred to texts that were literally printed on the photos.

Kappa statistics for the Instagram training set and the second Pinterest training set

Table B1. Kappa Statistics - Instagram Training Set

| <b>Column</b> | <b>Category</b>                                | <b>Kappa</b>                      |
|---------------|--|-----------------------------------|
| <b>1-4</b>    | N/A, contain Picture IDs and links to pictures |                                   |
| <b>5</b>      | Images without words                           | 0.701                             |
| <b>6</b>      | Images with words                              | 0.703                             |
| <b>7</b>      | English Language                               | 0.774                             |
| <b>8</b>      | Spanish/Portuguese Language                    | 0.77                              |
| <b>9</b>      | Other Language                                 | 0.626                             |
| <b>10</b>     | Not Relevant                                   | 0.823                             |
| <b>11</b>     | Relevant                                       | 0.819                             |
| <b>12</b>     | Symptoms                                       | 0.8                               |
| <b>13</b>     | Transmission                                   | 0.754                             |
| <b>14</b>     | Birth defects                                  | 0.817                             |
| <b>15</b>     | Treatment/Medical                              | 0.266 *                           |
| <b>16</b>     | Pregnancy                                      | 0.63                              |
| <b>17</b>     | Active Organizations                           | 0.681                             |
| <b>18</b>     | Laboratory                                     | 0.1 *                             |
| <b>19</b>     | Spread of virus                                | 0.506                             |
| <b>20</b>     | Prevention                                     | 0.504                             |
| <b>21</b>     | New cases                                      | 0.61                              |
| <b>22</b>     | Deaths   | N/A – no images<br>coded as death |
| <b>23</b>     | Travel   | 0.416                             |
| <b>24</b>     | Miscellaneous                                  | 0.3                               |

N/A: Not Applicable

\*Note: There were very few photos that match this category.

Table B2. Kappa Statistics - Pinterest Second Training Set

| <b>Column</b> | <b>Category</b>             | <b>Kappa</b> |
|---------------|-----------------------------|--------------|
| <b>1</b>      | Image without words         | 0.711        |
| <b>2</b>      | Image with words            | 0.711        |
| <b>3</b>      | English language            | 0.659        |
| <b>4</b>      | Spanish/Portuguese language | 0.728        |
| <b>5</b>      | Other language              | 0.568        |
| <b>6</b>      | Not Relevant                | 0.904        |
| <b>7</b>      | Relevant                    | 0.896        |
| <b>8</b>      | Symptoms                    | 0.871        |
| <b>9</b>      | Transmission                | 0.805        |
| <b>10</b>     | Birth defects               | 0.863        |
| <b>11</b>     | Treatment/Medical:          | 0.678*       |
| <b>12</b>     | Pregnancy                   | 0.865        |
| <b>13</b>     | Active Organizations        | 0.882        |
| <b>14</b>     | Laboratory                  | 0.723        |
| <b>15</b>     | Spread of virus             | 0.615        |
| <b>16</b>     | Prevention                  | 0.788        |
| <b>17</b>     | New cases                   | 0.572        |
| <b>18</b>     | Deaths                      | 0.665        |
| <b>19</b>     | Travel                      | 0.761        |
| <b>20</b>     | Miscellaneous               | 0.536        |

\*Note: There were very few photos that match this category.

Kappa statistics for research datasets

Table B3. Kappa Statistics - Instagram Dataset (Research)

| <b>Column</b> | <b>Category</b>                                | <b>Kappa</b> |
|---------------|--|--------------|
| <b>1-4</b>    | N/A, contain Picture IDs and links to pictures |              |
| <b>5</b>      | Images without words                           | 1.0          |
| <b>6</b>      | Images with words                              | 1.0          |
| <b>7</b>      | English Language                               | 0.993        |
| <b>8</b>      | Spanish/Portuguese Language                    | 0.984        |
| <b>9</b>      | Other Language                                 | 1.0          |
| <b>10</b>     | Not Relevant                                   | 0.985        |
| <b>11</b>     | Relevant                                       | 0.98         |
| <b>12</b>     | Symptoms                                       | 0.785        |
| <b>13</b>     | Transmission                                   | 0.873        |
| <b>14</b>     | Birth defects                                  | 0.88         |
| <b>15</b>     | Treatment/Medical                              | 0.947        |
| <b>16</b>     | Pregnancy                                      | 0.863        |
| <b>17</b>     | Active Organizations                           | 0.681        |
| <b>18</b>     | Laboratory                                     | 0.821        |
| <b>19</b>     | Spread of virus                                | 0.799        |
| <b>20</b>     | Prevention                                     | 0.982        |
| <b>21</b>     | New cases                                      | 1.0          |
| <b>22</b>     | Deaths   | 0.783        |
| <b>23</b>     | Travel   | 0.822        |
| <b>24</b>     | Miscellaneous                                  | 0.951        |

N/A: Not Applicable

Table B4. Kappa Statistics - Pinterest Dataset (Research)

| <b>Column</b> | <b>Category</b>             | <b>Kappa</b>                      |
|---------------|-----------------------------|-----------------------------------|
| <b>1</b>      | Image without words         | 0.992                             |
| <b>2</b>      | Image with words            | 0.969                             |
| <b>3</b>      | English language            | 0.898                             |
| <b>4</b>      | Spanish/Portuguese language | 0.735                             |
| <b>5</b>      | Other language              | 1.0                               |
| <b>6</b>      | Not Relevant                | 0.873                             |
| <b>7</b>      | Relevant                    | 0.856                             |
| <b>8</b>      | Symptoms                    | 0.871                             |
| <b>9</b>      | Transmission                | 0.981                             |
| <b>10</b>     | Birth defects               | 0.795                             |
| <b>11</b>     | Treatment/Medical           | 0.948                             |
| <b>12</b>     | Pregnancy                   | 0.958                             |
| <b>13</b>     | Active Organizations        | 0.883                             |
| <b>14</b>     | Laboratory                  | 0.811                             |
| <b>15</b>     | Spread of virus             | 0.855                             |
| <b>16</b>     | Prevention                  | 0.956                             |
| <b>17</b>     | New cases                   | 0.856                             |
| <b>18</b>     | Deaths                      | N/A – no images<br>coded as death |
| <b>19</b>     | Travel                      | 0.783                             |
| <b>20</b>     | Miscellaneous               | 0.865                             |

N/A: Not Applicable

## Results

Table B5. Distribution of language of embedded words in relevant photos from Pinterest and Instagram.

|                              | <b>Pinterest (% , N=164)</b> | <b>Instagram (% , N=144)</b> |
|------------------------------|------------------------------|------------------------------|
| <b>English</b>               | 151 (92)                     | 66 (46)                      |
| <b>Spanish or Portuguese</b> | 14 (9)                       | 77 (53)                      |
| <b>Other</b>                 | 2 (1)                        | 0 (0)                        |

Fisher's exact test, p-value < 2.2e-16.

### Pairwise odds ratio of co-occurrence of topics

We used univariate logistic regression to calculate the odds ratio of co-occurrence of two content categories among ZIKV-relevant Instagram and Pinterest photos.

Table B6 presents an example. Among ZIKV-related Instagram photos, there were seven that were coded as both “birth defect” and “treatment,” and 119 that were neither “birth defect” nor “treatment.” Nine were coded as “yes” for “treatment” and “no” for “birth defect,” while another nine were coded as “no” for “treatment” and “yes” for “birth defect.”

The odds ratio is the ratio between the odds of coding as “birth defect” among photos coded “yes” for “treatment” (7/9) and the odds of coding as “birth defect” among photos coded “no” for “treatment” (9/119), i.e. odds ratio = (7/9) / (9/119) = 10.284. Therefore, for a ZIKV-related Instagram photo that is coded as “treatment”, it is ten times more likely to be co-coded as “birth defect,” as compared to ZIKV-related Instagram photos that is not coded for “treatment.”

Table B6. Odds ratio calculation for co-occurrence of treatment category and birth defect category among ZIKV-related Instagram photos.

|                           | <b>Treatment “No”</b> | <b>Treatment “Yes”</b> |
|---------------------------|-----------------------|------------------------|
| <b>Birth defect “No”</b>  | 119                   | 9                      |
| <b>Birth defect “Yes”</b> | 9                     | 7                      |

Tables B7 and B8 present pairwise odds ratios for ZIKV-related Instagram and Pinterest photos respectively. We excluded the “Miscellaneous” category because by definition, a ZIKV-related photo was coded as “Miscellaneous” if it was not coded for any other category.

*Instagram*

Table B7. The odds ratio (p-value in parentheses) for co-occurrence of two content categories of ZIKV-related Instagram photos obtained by pairwise univariate logistic regression.

|    | 3 | 4  | 5              | 6                 | 7               | 8  | 9              | 10 | 11              | 13              | 14             | 15              |
|----|---|----|----------------|-------------------|-----------------|----|----------------|----|-----------------|-----------------|----------------|-----------------|
| 3  | - | NS | 4.52<br>(0.01) | NS                | NS              | NS | NS             | NS | NS              | NS              | NS             | NS              |
| 4  |   | -  | NS             | NS                | NS              | NS | 0.11<br>(0.04) | NS | NS              | NS              | NS             | NS              |
| 5  |   |    | -              | 10.28<br>(0.0001) | 5.74<br>(0.002) | NS | NS             | NS | NS              | NS              | NS             | 6.78<br>(0.007) |
| 6  |   |    |                | -                 | 5.74<br>(0.002) | NS | NS             | NS | NS              | NS              | 9.00<br>(0.03) | NS              |
| 7  |   |    |                |                   | -               | NS | NS             | NS | 4.57<br>(0.001) | NS              | NS             | 8.07<br>(0.002) |
| 8  |   |    |                |                   |                 | -  | NS             | NS | NS              | NS              | NS             | NS              |
| 9  |   |    |                |                   |                 |    | -              | NS | NS              | 19.29<br>(0.04) | NS             | NS              |
| 10 |   |    |                |                   |                 |    |                | -  | NS              | NS              | NS             | NS              |
| 11 |   |    |                |                   |                 |    |                |    | -               | NS              | NS             | NS              |
| 13 |   |    |                |                   |                 |    |                |    |                 | -               | NS             | NS              |
| 14 |   |    |                |                   |                 |    |                |    |                 |                 | -              | NS              |
| 15 |   |    |                |                   |                 |    |                |    |                 |                 |                | -               |

NS: Not significant

Categories:

3. Symptoms; 4. Transmission; 5. Birth defects; 6. Treatment/Medical; 7. Pregnancy; 8. Active Organizations; 9. Laboratory; 10. Spread of virus; 11. Prevention; 13. New cases; 14. Deaths; 15. Travel.

Note: Category 12 was removed from the codebook during the pilot phase of this study.

*Pinterest*

Table B8. The odds ratio (p-value in parentheses) for co-occurrence of two content categories of ZIKV-related Pinterest photos obtained by pairwise univariate logistic regression.

|    | 3 | 4  | 5  | 6               | 7              | 8  | 9  | 10              | 11              | 13 | 14 | 15               |
|----|---|----|----|-----------------|----------------|----|----|-----------------|-----------------|----|----|------------------|
| 3  | - | NS | NS | 3.65<br>(0.006) | NS             | NS | NS | NS              | 3.02<br>(0.005) | NS | NA | NS               |
| 4  |   | -  | NS | NS              | NS             | NS | NS | NS              | 2.55<br>(0.009) | NS | NA | NS               |
| 5  |   |    | -  | NS              | 2.63<br>(0.03) | NS | NS | NS              | NS              | NS | NA | NS               |
| 6  |   |    |    | -               | 3.27<br>(0.02) | NS | NS | NS              | 3.00<br>(0.02)  | NS | NA | NS               |
| 7  |   |    |    |                 | -              | NS | NS | 3.49<br>(0.005) | NS              | NS | NA | 3.00<br>(0.02)   |
| 8  |   |    |    |                 |                | -  | NS | NS              | NS              | NS | NA | 8.99<br>(<0.001) |
| 9  |   |    |    |                 |                |    | -  | NS              | NS              | NS | NA | NS               |
| 10 |   |    |    |                 |                |    |    | -               | NS              | NS | NA | NS               |
| 11 |   |    |    |                 |                |    |    |                 | -               | NS | NA | NS               |
| 13 |   |    |    |                 |                |    |    |                 |                 | -  | NA | NS               |
| 14 |   |    |    |                 |                |    |    |                 |                 |    | -  | NS               |
| 15 |   |    |    |                 |                |    |    |                 |                 |    |    | -                |

NS: Not significant

NA: Applicable

Categories:

3. Symptoms; 4. Transmission; 5. Birth defects; 6. Treatment/Medical; 7. Pregnancy; 8. Active Organizations; 9. Laboratory; 10. Spread of virus; 11. Prevention; 13. New cases; 14. Deaths; 15. Travel.

Note: Category 12 was removed from the codebook during the pilot phase of this study. There were no ZIKV-related Pinterest photos that were coded as Category 14 (Deaths).



## References

1. Coles P. Full Page Screen Capture. Available from: <https://chrome.google.com/webstore/detail/full-page-screen-capture/fdpohaocaechifmmbbbbknolclacl?hl=en-US>. Accessed April 23, 2016