Incident bivariate “Chi test”

|  |  |  |
| --- | --- | --- |
| **Variable Name** | **Pearson chi2** | **P Value** |
| Age Category | 2.7848 | 0.248 |
| Sex | 0.0210 | 0.885 |
| Child Education Level | 10.5493 | 0.005 |
| Religion | 3.8380 | 0.147 |
| BMI Category | 5.2882 | 0.152 |
| Child Tribe | 2.8586 | 0.239 |
| Orphanhood | 1.4443 | 0.486 |
| Socioeconomic status | 6.9153 | 0.009 |
| Support Program Enrolment | 1.7270 | 0.189 |
| Child ART Status | 0.3432 | 0.558 |
| Premature Birth | 2.6265 | 0.269 |
| Born with HIV/AIDS | 1.9384 | 0.379 |
| Stress | 24.9421 | 0.000 |
| Worst HIV Clinical Stage | 1.2927 | 0.731 |
| CD4 Category | 0.6220 | 0.733 |
| Viral load | 3.6845 | 0.158 |
| stin2vntr | 0.4059 | 0.816 |
| httlpr1 | 1.1331 | 0.567 |
| HTTLPRrs35531 | 7.6484 | 0.177 |
| rs35531 | 7.0999 | 0.029 |
| Rs10482605 | 1.5340 | 0.464 |
| Rs1360780 | 4.7025 | 0.095 |
| rs1386494 | 0.8901 | 0.641 |
| rs1843809 | 1.3920 | 0.499 |
| rs34517220 | 6.7940 | 0.033 |

**For Incident IMD**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **STATA** | **Machine Learning** | | **Agreement** |
| **Variable/Feature** | **P Value** | **Feature** | **Importance** |  |
| Baseline telomere length | n/a |  | 0.121 | n/a |
| Telomere\_length\_12Months | n/a |  | 0.099 | n/a |
| Telomere Length Difference | n/a |  | 0.096 | n/a |
| Stress | 0.000 | Stress | 0.077 | Yes |
| Child Education | 0.005 | Orphanhood | 0.031 | No |
| Socioeconomic status | 0.009 | Socioeconomic status | 0.030 | Yes |
| rs35531 | 0.029 | rs35531 | 0.034 | Yes |
| rs34517220 | 0.033 | rs34517220 | 0.038 | Yes |
| Rs1360780 | 0.095 | Rs1360780 | 0.034 | Yes |
| Religion | 0.147 | Worst\_HIV\_Stage | 0.040 | No |
| BMI Category | 0.152 | Age\_Category | 0.034 | No |
| Viral load | 0.158 | Viral load | 0.045 | Yes |
| HTTLPRrs35531 | 0.177 | HTTLPRrs35531 | 0.045 | Yes |
| **Total Agreement** |  |  |  | =7/10 = **70%** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **STATA** | **Machine Learning** | | **Agreement** |
| **Variable/Feature** | **P Value** | **Feature** | **Importance** |  |
| rs35531 | 0.029 | rs35531 | 0.034 | Yes |
| Rs1360780 | 0.095 | Rs1360780 | 0.034 | Yes |
| HTTLPRrs35531 | 0.177 | HTTLPRrs35531 | 0.045 | Yes |
| **Total Agreement** |  |  |  | =7/10 = **70%** |

|  |  |
| --- | --- |
| Feature | Importance |
| Baseline\_telomere\_length |  |
| Telomere\_length\_12Months |  |
| Telomere\_Lenth\_Difference |  |
| Stress | 0.077 |
| Viralload\_Category | 0.045 |
| HTTLPRrs35531 | 0.045 |
| Age\_Category | 0.040 |
| rs34517220\_polymorphism | 0.038 |
| Worst\_HIV\_Stage | 0.034 |
| rs35531\_polymorphism | 0.034 |
| Rs1360780\_polymorphism | 0.034 |
| orphanhood | 0.031 |
| Socio\_Economic\_Status | 0.030 |
| rs1843809\_polymorphism | 0.028 |
| httlpr1\_polymorphism | 0.028 |
| stin2vntr\_ | 0.024 |
| Gender\_of\_Child | 0.023 |
| Childs\_Education\_Level | 0.023 |
| rs1386494\_polymorphism | 0.022 |
| Tribe\_of\_Child | 0.021 |
| CD4\_category | 0.021 |
| Rs10482605\_polymorphism | 0.019 |
| Religion | 0.019 |

DISCUSSION

|  |  |
| --- | --- |
| Worst\_HIV\_Stage | World Health Organization (WHO) clinical staging of HIV disease have a direct impact on emotional well-being, as progression to advanced clinical stages (e.g., WHO Stage III or IV) may involve a higher burden of physical symptoms and complications, which can cause anxiety and depression (Bartlett et al., 2019). |
| Viral load | High viral loads indicate active viral replication and uncontrolled HIV infection, which can lead to persistent stress and anxiety (Zielińska-Wieniawska et al., 2021b). |
| Socioeconomic status | Low family income and living in poverty are associated with increased stress, anxiety and depression (Lwidiko et al., 2018). |
| Orphaned | Heightened psychological distress and a lack of emotional support (Jorns-Presentati et al., 2021). |
| Telomere Length | Significantly shorter in patients with mental disorders like depression, anxiety, and stress compared to individuals without these disorders (Kalungi, Womersley, et al., 2019). One reason is biological aging. |
| Stress | Psychological stress has been linked to higher oxidative stress, lower telomerase activity, and shortened telomere length, which may contribute to the biological aging observed in individuals with mental disorders like depression, anxiety, and stress[3](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5778888/)[4](https://www.medrxiv.org/content/10.1101/2022.03.29.22273140v1). |
| Age | Older age has been linked to shorter telomeres in individuals with mental disorders like depression and anxiety[3](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5778888/). |
| rs34517220 (*TPH2*) | The tryptophan hydroxylase gene (*TPH*) has also been reported to be associated with IMDs, as this gene encodes Tryptophan hydroxylase enzyme, which is responsible for the synthesis of serotonin, a neurotransmitter that plays a crucial role in mood regulation (G. L. Chen & Miller, 2013).  *rs34517220* SNP has been reported to modulate *TPH2* expression by altering transcription factor binding sites (Womersley, n.d.). |
| Rs1360780 (*FKBP5* Gene) | The *FKBP5* gene, also known as *FKBP* *prolyl* isomerase 5, is a stress-inducible gene that encodes *FK506*-binding protein 51 (*FKBP51*) (Wang et al., 2018). This gene is highly expressed in brain regions associated with stress and anxiety responses. Genetic variations of *FKBP5* have been linked to the pathogenesis of psychiatric disorders, increasing vulnerability to stress-related anxiety and depression among adolescents and children with HIV (Owalla et al., 2021). Certain genetic variants of *FKBP5* such single-nucleotide polymorphisms within the *FK506* binding protein 5 (*rs1360780*) and glucocorticoid receptor (*rs10482605*) genes may lead to a heightened response to stress, potentially increasing the risk of developing IMDs, such as depression and anxiety (Owalla et al., 2021). |
| HTTLPRrs35531 and rs35531 ( *SLC6A4*) | The central nervous system relies heavily on serotonergic pathways to regulate mood and anxiety. This intricate system involves the serotonin transporter gene (*SLC6A4*), which encodes the serotonin transporter (*5-HTT*), a critical player in controlling serotonin levels within the synaptic clefts (Matsumoto et al., 2014). The *5-HTT* is a target for selective serotonin re-uptake inhibitors, drugs that enhance serotonin action at synapses (Matsumoto et al., 2014).  The *rs25531* SNP, located proximally to 5-HTTLPR, influences *SLC6A4* expression, creating functional AP2 transcription-factor binding sites (Wankerl et al., 2014). Acute stress and chronic stress have been reported to interact with *5-HTTLPR-rs25531* haplotypes to influence the risk for depression and anxiety among Ugandan children and adolescents living with HIV (Kalungi et al., 2021). |

|  |  |
| --- | --- |
| **Variable** | **Association with Mental Disorders** |
| Worst HIV Stage | - Advanced clinical stages (e.g., WHO Stage III or IV) may involve a higher burden of physical symptoms and complications, causing anxiety and depression (Bartlett et al., 2019). |
| Viral Load | - High viral loads indicate uncontrolled HIV infection, leading to persistent stress and anxiety (Zielińska-Wieniawska et al., 2021b). |
| Socioeconomic Status | - Low SES are associated with increased stress, anxiety and depression (Lwidiko et al., 2018). |
| Orphanhood | - Heightened psychological distress and a lack of emotional support |
| Telomere Length | - Significantly shorter in patients with mental disorders (Kalungi, Womersley, et al., 2019). |
| Stress | - Psychological stress is linked to higher oxidative stress, lower telomerase activity, and shortened telomere length, contributing to biological aging in individuals with mental disorders |
| Age | - Older age is associated with shorter telomeres in individuals with mental disorders |
| *rs34517220* (*TPH2*) | - The *tryptophan hydroxylase* gene (TPH) is associated with mental disorders, as it encodes an enzyme responsible for serotonin synthesis, crucial for mood regulation anxiety (Owalla et al., 2021). |
| *Rs1360780* (*FKBP5* Gene) | - Linked to the pathogenesis of psychiatric disorders, increasing vulnerability to stress-related anxiety and depression |
| *HTTLPRrs35531* and *rs35531* (*SLC6A4*) | - The serotonin transporter gene (*SLC6A4*) is critical for controlling serotonin levels and regulating mood and anxiety |

PSYCHOSOCIAL FACTORS

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classifier | F1 Score | Precision | Recall | Accuracy | ROC AUC |
| Logistic Regression | 0.665 | 0.691 | 0.697 | 0.697 | 0.691 |
| Gaussian Naive Bayes | 0.541 | 0.616 | 0.531 | 0.531 | 0.618 |
| K-Nearest Neighbors | 0.626 | 0.627 | 0.632 | 0.632 | 0.618 |
| Decision Tree | 0.553 | 0.570 | 0.550 | 0.567 | 0.553 |
| Support Vector Classifier | 0.594 | 0.627 | 0.673 | 0.673 | 0.690 |
| Random Forest | 0.621 | 0.635 | 0.636 | 0.625 | 0.612 |
| Gradient Boosting | 0.656 | 0.662 | 0.683 | 0.683 | 0.670 |
| AdaBoost | 0.667 | 0.679 | 0.693 | 0.693 | 0.659 |
| Linear Discriminant Analysis | 0.658 | 0.681 | 0.686 | 0.686 | 0.693 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **STATA** | | **Machine Learning** | | **Agreement** |
| Variable/Feature | P Value | Feature | Importance |  |
| Stress | 0.000 | Stress | 0.077 | Yes |
| Socioeconomic status | 0.009 | Socioeconomic status | 0.030 | Yes |

Biological factors Incident

Spots checking

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classifier | F1 Score | Precision | Recall | Accuracy | ROC AUC |
| KNeighbors Classifier | 0.635 | 0.641 | 0.667 | 0.667 | 0.578 |
| Decision Tree Classifier | 0.625 | 0.614 | 0.674 | 0.674 | 0.596 |
| Random Forest Classifier | 0.624 | 0.624 | 0.681 | 0.694 | 0.597 |
| Gradient Boosting Classifier | 0.609 | 0.611 | 0.677 | 0.677 | 0.559 |
| GaussianNB | 0.600 | 0.614 | 0.656 | 0.656 | 0.557 |
| Linear Discriminant Analysis | 0.565 | 0.568 | 0.677 | 0.677 | 0.568 |
| Logistic Regression | 0.554 | 0.521 | 0.674 | 0.674 | 0.571 |
| AdaBoost Classifier | 0.553 | 0.515 | 0.663 | 0.663 | 0.542 |
| SVC | 0.552 | 0.519 | 0.670 | 0.670 | 0.533 |

Gentic prectiords incident

“

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classifier | F1 Score | Precision | Recall | Accuracy | ROC AUC |
| AdaBoostClassifier | 0.598 | 0.607 | 0.653 | 0.653 | 0.591 |
| GaussianNB | 0.589 | 0.589 | 0.650 | 0.650 | 0.552 |
| KNeighborsClassifier | 0.582 | 0.573 | 0.615 | 0.615 | 0.523 |
| GradientBoostingClassifier | 0.580 | 0.586 | 0.626 | 0.626 | 0.496 |
| LinearDiscriminantAnalysis | 0.562 | 0.553 | 0.639 | 0.639 | 0.538 |
| LogisticRegression | 0.554 | 0.538 | 0.650 | 0.650 | 0.545 |
| SVC | 0.536 | 0.482 | 0.650 | 0.650 | 0.429 |
| RandomForestClassifier | 0.532 | 0.505 | 0.561 | 0.547 | 0.472 |
| DecisionTreeClassifier | 0.518 | 0.512 | 0.520 | 0.513 | 0.470 |

PERSISTENT IMD

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classifier | F1 Score | Precision | Recall | Accuracy | ROC AUC |
| KNeighborsClassifier | 0.787 | 0.761 | 0.827 | 0.827 | 0.595 |
| GradientBoostingClassifier | 0.782 | 0.763 | 0.833 | 0.837 | 0.673 |
| SVC | 0.777 | 0.718 | 0.847 | 0.847 | 0.637 |
| RandomForestClassifier | 0.774 | 0.717 | 0.837 | 0.840 | 0.677 |
| AdaBoostClassifier | 0.767 | 0.746 | 0.793 | 0.793 | 0.592 |
| LogisticRegression | 0.765 | 0.715 | 0.823 | 0.823 | 0.643 |
| LinearDiscriminantAnalysis | 0.754 | 0.712 | 0.803 | 0.803 | 0.610 |
| DecisionTreeClassifier | 0.747 | 0.755 | 0.748 | 0.731 | 0.517 |
| GaussianNB | 0.661 | 0.738 | 0.655 | 0.655 | 0.555 |

HYPER PARA

|  |  |  |
| --- | --- | --- |
| Classifier | Accuracy (Original) | Accuracy (Updated) |
| Logistic Regression | 0.917 | 0.847 |
| Support Vector Classifier | 0.917 | 0.847 |
| Gradient Boosting | 0.917 | 0.847 |
| AdaBoost | 0.917 | 0.847 |
| Random Forest | 0.915 | 0.844 |
| K-Nearest Neighbors | 0.912 | 0.840 |
| Linear Discriminant Analysis | 0.890 | 0.803 |
| Decision Tree | 0.851 | 0.751 |
| Gaussian Naive Bayes | 0.747 | 0.655 |

Comapred to Statistics

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **STATA** | **Machine Learning** | | **Agreement** |
| **Variable/Feature** | **P Value** | **Feature** | **Importance** |  |
| Stress | P<0.05 | Stress | 0.077 | Yes |
| Age category | P<0.05 | Age category | 0.031 | Yes |
| Education Level | P<0.05 |  | 0.038 | No |
| BMI Category | P<0.05 | BMI Category | 0.034 | Yes |
|  |  | HTTLPRrs35531 | 0.045 | No |
| **Total Agreement** |  |  |  | =3/5 = 6**0%** |