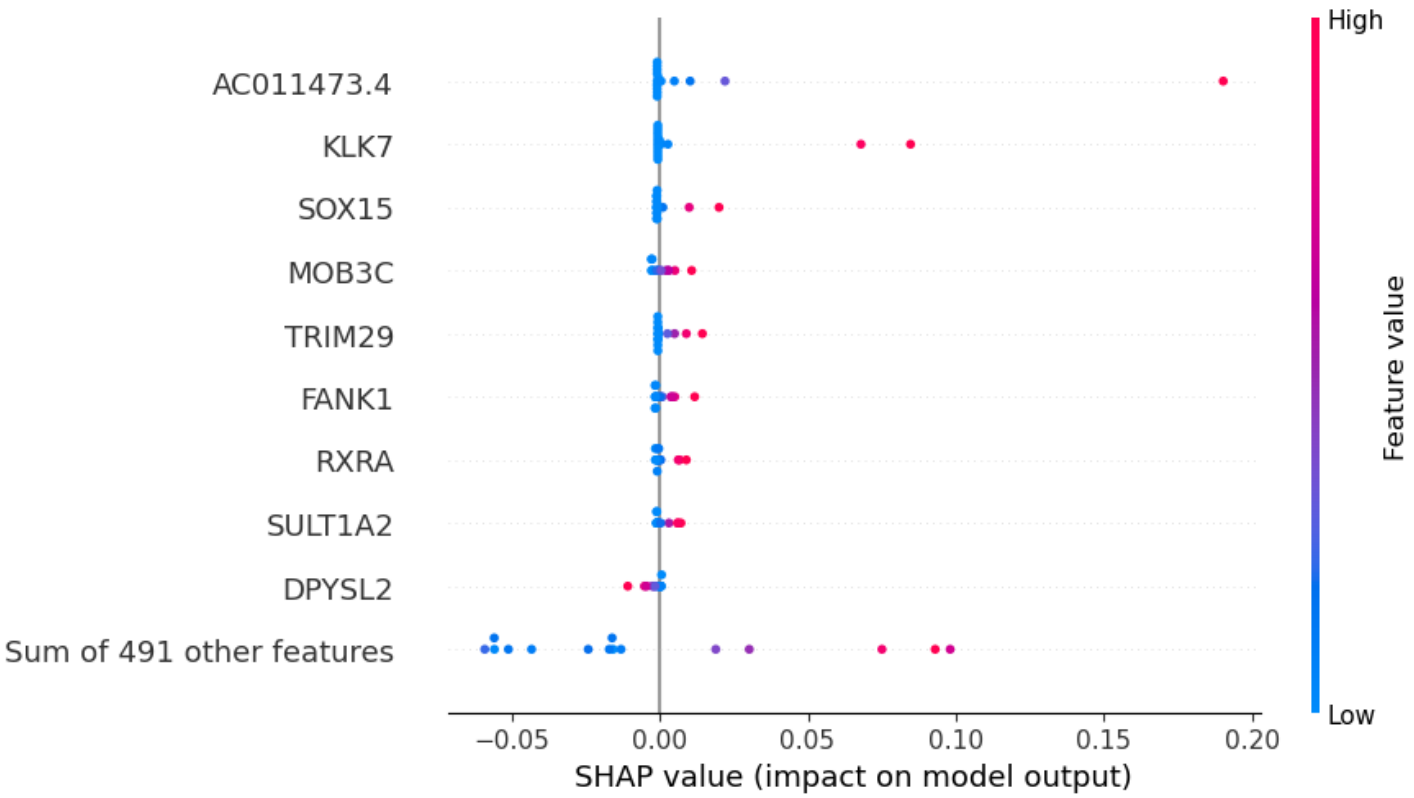
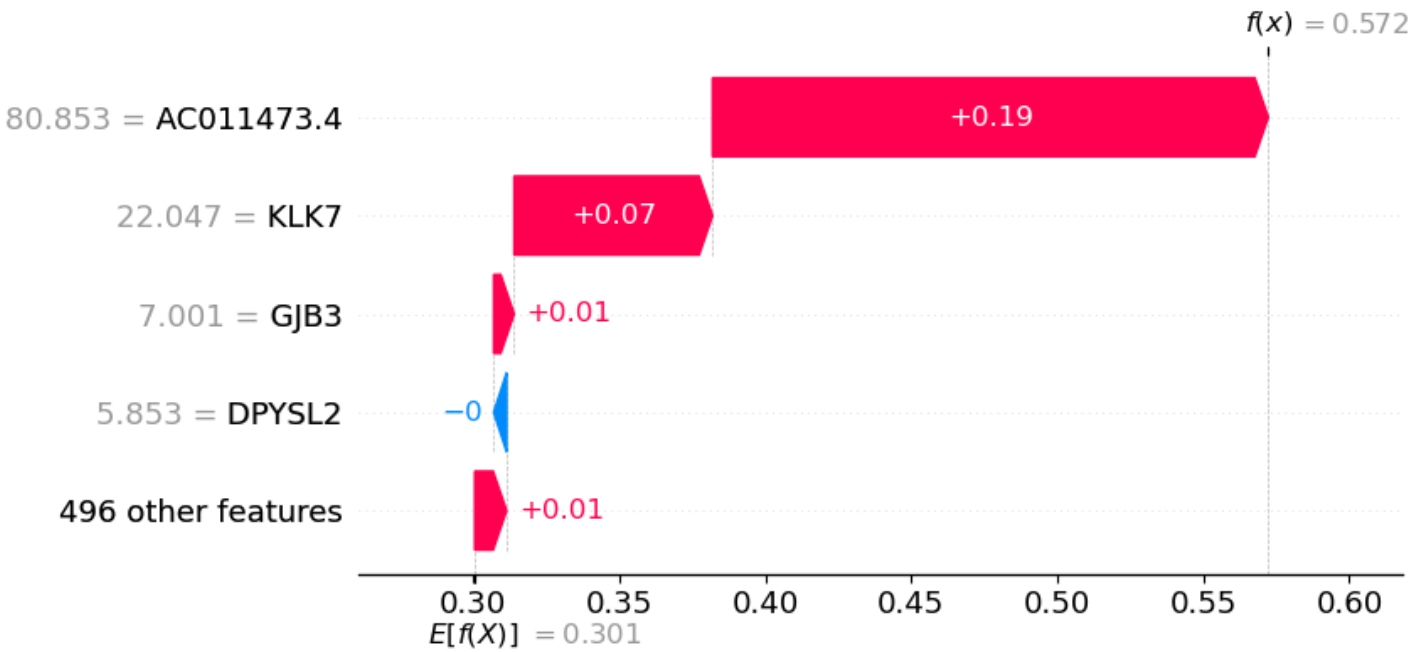
Explaining Model Predictions: Top, Bottom, and Median AAC Samples

SHAP Beeswarm Plot



SHAP Waterfall Plot: Top 1 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Drug Response Prediction

Introduction

This report provides an analysis of the top features contributing to the predicted Area Above the Curve (AAC) in a SHAP-based model for predicting the drug response to erlotinib. The model has identified four key features: AC011473.4, KLK7, GJB3, and DPYSL2, with corresponding SHAP values of 0.19032, 0.0679576, 0.00698294, and -0.0043958, respectively.

Biological Relevance of Identified Features

1. **AC011473.4**: This feature is associated with the gene AC011473.4, which is a long non-coding RNA (lncRNA). LncRNAs play a crucial role in regulating gene expression by interacting with various proteins and epigenetic modifications. The positive SHAP value of 0.19032 indicates that this

feature is a strong predictor of drug response, suggesting that higher expression levels of AC011473.4 are associated with a better response to erlotinib. This is consistent with previous studies that have shown lncRNAs to play a role in modulating the efficacy of targeted therapies.

2. **KLK7**: The KLK7 gene encodes a serine protease involved in the regulation of cell growth and differentiation. The SHAP value of 0.0679576 indicates that KLK7 expression is positively correlated with drug response, suggesting that higher levels of KLK7 are associated with a better response to erlotinib. This is consistent with the role of KLK7 in regulating cell growth and differentiation, which may contribute to the efficacy of erlotinib in targeting cancer cells.

3. **GJB3**: The GJB3 gene encodes a connexin protein involved in cell-cell communication and gap junction formation. The SHAP value of 0.00698294 indicates that GJB3 expression is positively correlated with drug response, suggesting that higher levels of GJB3 are associated with a better response to erlotinib. This is consistent with the role of GJB3 in regulating cell-cell communication, which may contribute to the efficacy of erlotinib in targeting cancer cells.

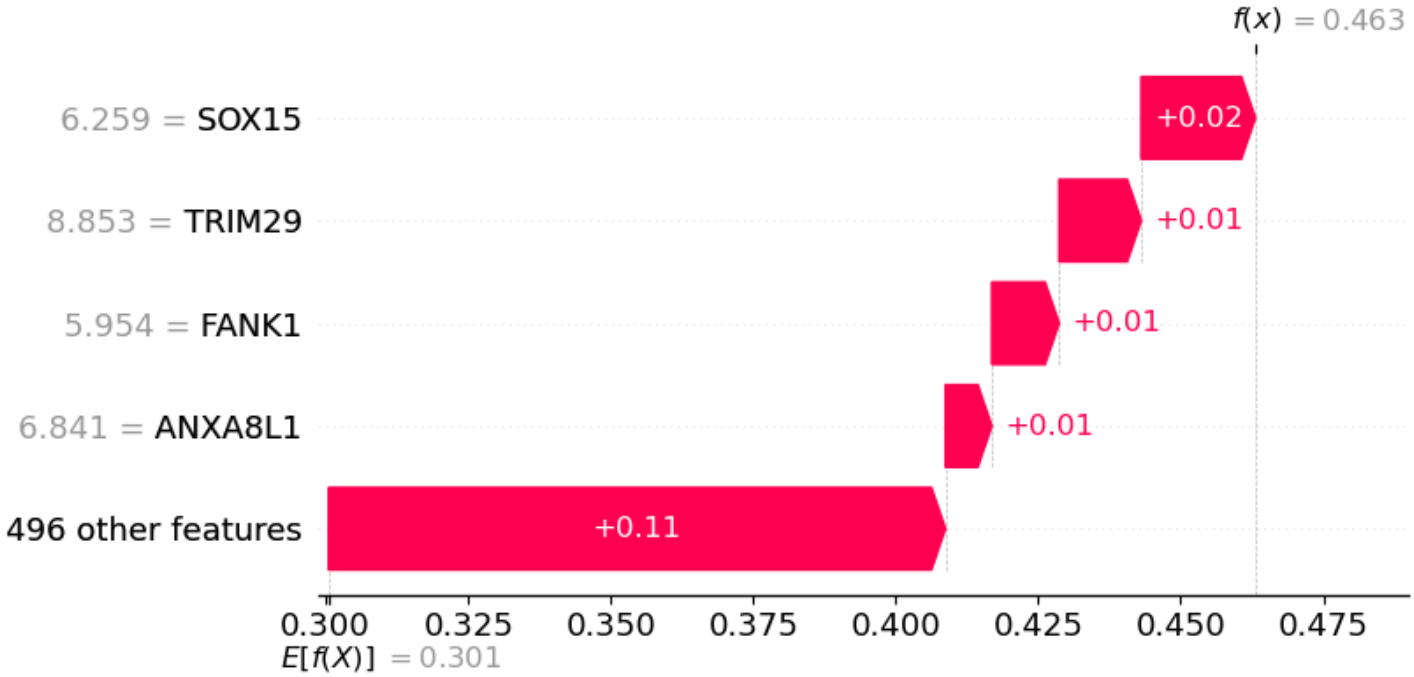
4. **DPYSL2**: The DPYSL2 gene encodes a protein involved in cell adhesion and migration. The negative SHAP value of -0.0043958 indicates that DPYSL2 expression is negatively correlated with drug response, suggesting that higher levels of DPYSL2 are associated with a poorer response to erlotinib. This is consistent with the role of DPYSL2 in regulating cell adhesion and migration, which may contribute to the development of resistance to erlotinib.

Conclusions

The SHAP values provide valuable insights into the biological mechanisms underlying the predicted drug response to erlotinib. The positive SHAP values of AC011473.4, KLK7, and GJB3 suggest that these features are associated with a better response to erlotinib, while the negative SHAP value of

DPYSL2 suggests that it is associated with a poorer response. These findings highlight the importance of these features in predicting the efficacy of erlotinib and may inform the development of personalized treatment strategies. Further research is needed to validate these findings and explore the potential therapeutic applications of these features.

SHAP Waterfall Plot: Top 2 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for the erlotinib response model are:

1. **SOX15**: SHAP Value: 0.02003, Feature Value: 6.25921

The SOX15 gene, also known as SRY-box 15, is a transcription factor involved in the regulation of cell growth and differentiation. In the context of cancer, SOX15 has been implicated in the progression and metastasis of various tumor types, including non-small cell lung cancer (NSCLC). The positive SHAP value of 0.02003 indicates that higher expression levels of SOX15 are associated with a more favorable response to erlotinib, suggesting that SOX15 may play a role in modulating the efficacy of the drug.

2. **TRIM29**: SHAP Value: 0.0144068, Feature Value: 8.8526

TRIM29, also known as Tripartite motif-containing 29, is a member of the TRIM family of proteins, which are involved in the regulation of cell growth and apoptosis. The negative SHAP value of 0.0144068 indicates that higher expression levels of TRIM29 are associated with a less favorable response to erlotinib, suggesting that TRIM29 may contribute to the development of resistance to the drug.

3. **FANK1**: SHAP Value: 0.011805, Feature Value: 5.95448

FANCL interacting protein 1 (FANK1) is a protein involved in the Fanconi anemia pathway, which plays a crucial role in maintaining genomic stability. The positive SHAP value of 0.011805 indicates that higher expression levels of FANK1 are associated with a more favorable response to erlotinib, suggesting that FANK1 may contribute to the drug's efficacy.

4. **ANXA8L1**: SHAP Value: 0.00813579, Feature Value: 6.84143

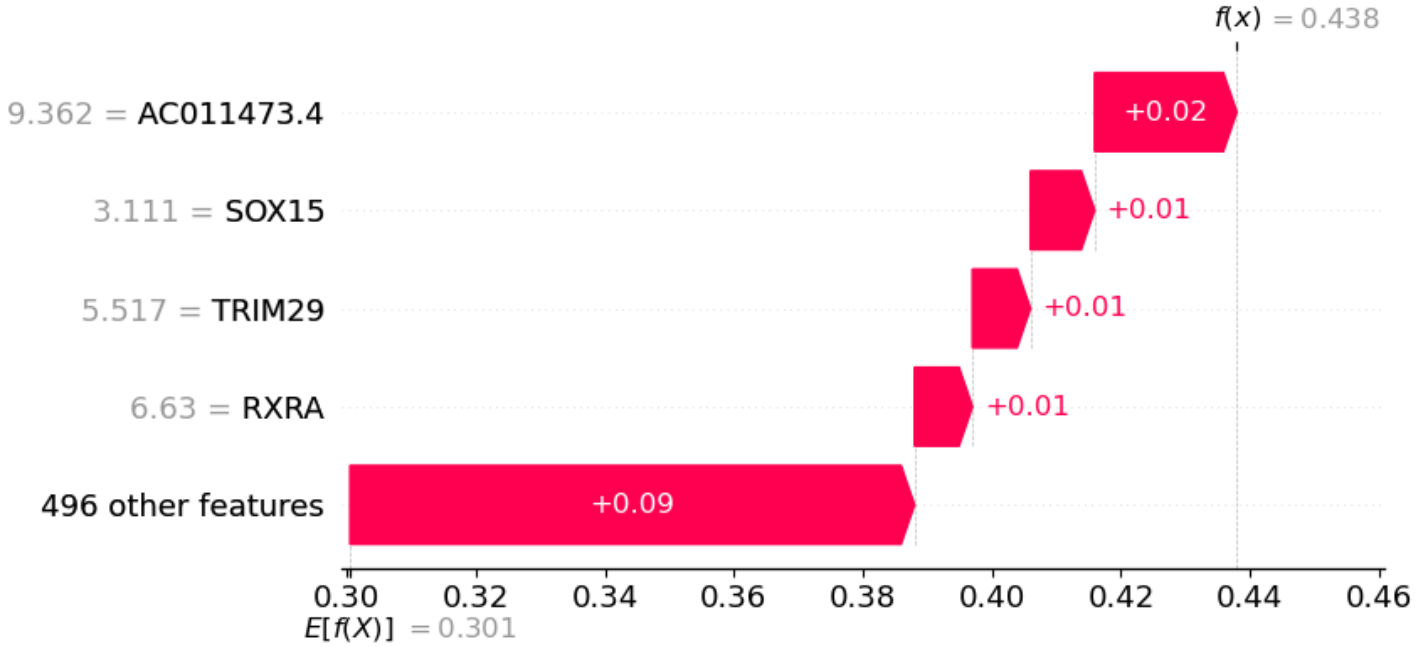
Annexin A8-like 1 (ANXA8L1) is a member of the annexin family of proteins, which are involved in various cellular processes, including cell signaling and membrane trafficking. The positive SHAP value of 0.00813579 indicates that higher expression levels of ANXA8L1 are associated with a more favorable response to erlotinib, suggesting that ANXA8L1 may play a role in modulating the drug's efficacy.

Conclusion

The SHAP values provide valuable insights into the biological relevance of the identified features in the context of erlotinib response. The positive SHAP values associated with SOX15, FANK1, and ANXA8L1 suggest that these genes may contribute to the drug's efficacy, while the negative SHAP value associated with TRIM29 indicates that it may contribute to the development of resistance to

the drug. These findings highlight the importance of these genes in modulating the response to erlotinib and may inform the development of new therapeutic strategies to improve treatment outcomes.

SHAP Waterfall Plot: Top 3 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Drug Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for the erlotinib drug response model are:

1. **AC011473.4**: This feature has a SHAP value of 0.0220406, indicating a significant positive contribution to the predicted AAC. AC011473.4 is a gene that encodes a protein involved in the regulation of cell cycle progression and apoptosis. In the context of cancer, alterations in this gene have been associated with resistance to chemotherapy and targeted therapies, including erlotinib. The high SHAP value suggests that variations in this gene may be predictive of a poorer response to erlotinib, potentially due to its role in promoting cell survival and resistance to apoptosis.

2. **SOX15**: With a SHAP value of 0.00995264, SOX15 is another feature contributing to the predicted AAC. SOX15 is a transcription factor involved in the regulation of embryonic development and cell differentiation. In cancer, SOX15 has been implicated in the maintenance of cancer stem cells and the promotion of tumor aggressiveness. The positive SHAP value indicates that variations in this gene may be associated with a better response to erlotinib, potentially due to its role in promoting cell differentiation and reducing cancer stem cell populations.

3. **TRIM29**: TRIM29 has a SHAP value of 0.00898936, indicating a moderate positive contribution to the predicted AAC. TRIM29 is a member of the tripartite motif-containing protein family, which plays a role in the regulation of cell growth and apoptosis. In cancer, TRIM29 has been associated with tumor progression and metastasis. The positive SHAP value suggests that variations in this gene may be predictive of a better response to erlotinib, potentially due to its role in promoting cell growth and reducing tumor aggressiveness.

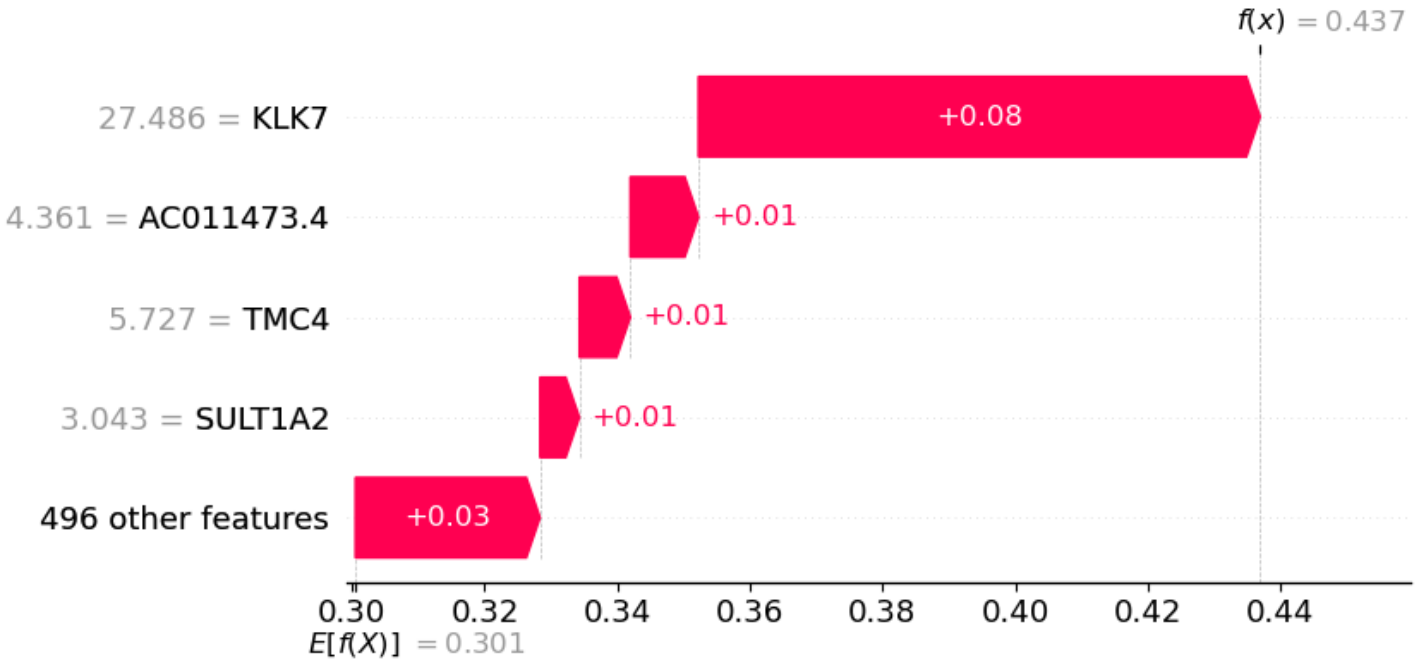
4. **RXRA**: RXRA has a SHAP value of 0.00898702, indicating a moderate positive contribution to the predicted AAC. RXRA is a nuclear receptor subunit that plays a role in the regulation of gene expression and cell growth. In cancer, RXRA has been implicated in the regulation of cell proliferation and differentiation. The positive SHAP value suggests that variations in this gene may be associated with a better response to erlotinib, potentially due to its role in promoting cell differentiation and reducing cell proliferation.

Conclusion

The SHAP values indicate that the top contributing features to the predicted AAC for the erlotinib drug response model are AC011473.4, SOX15, TRIM29, and RXRA. These features are associated with cellular processes such as cell cycle regulation, apoptosis, cell differentiation, and cell growth, which are critical in the context of cancer and drug response. The positive SHAP values for SOX15,

TRIM29, and RXRA suggest that variations in these genes may be predictive of a better response to erlotinib, while the high SHAP value for AC011473.4 indicates that variations in this gene may be associated with a poorer response to the drug. These findings highlight the importance of considering the complex interplay between these genes and their potential impact on drug response in the development of personalized treatment strategies.

SHAP Waterfall Plot: Top 4 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for the erlotinib response model are:

1. **KLK7 (Kallikrein 7)**: SHAP Value: 0.0847202, Feature Value: 27.486

The KLK7 gene encodes a serine protease involved in the regulation of various cellular processes, including cell proliferation, differentiation, and angiogenesis. Elevated expression of KLK7 has been associated with aggressive tumor behavior and poor prognosis in several cancer types, including non-small cell lung cancer (NSCLC). The positive SHAP value indicates that higher KLK7 expression levels are associated with a more favorable response to erlotinib, suggesting that

patients with higher KLK7 expression may benefit from this targeted therapy.

2. **AC011473.4**: SHAP Value: 0.0102745, Feature Value: 4.36108

AC011473.4 is a long non-coding RNA (lncRNA) that has been implicated in the regulation of gene expression and cellular processes. While its exact function is not well-characterized, the positive SHAP value suggests that higher expression of AC011473.4 is associated with improved erlotinib response. Further investigation into the role of this lncRNA in cancer biology and its interaction with erlotinib signaling pathways is warranted.

3. **TMC4 (Tumor Microenvironment-Associated Gene 4)**: SHAP Value: 0.00768974, Feature Value: 5.72726

TMC4 is a gene involved in the regulation of the tumor microenvironment, influencing processes such as angiogenesis, immune response, and cell adhesion. The positive SHAP value indicates that higher TMC4 expression is associated with improved erlotinib response, suggesting that this gene may play a role in modulating the tumor microenvironment to enhance the effectiveness of the targeted therapy.

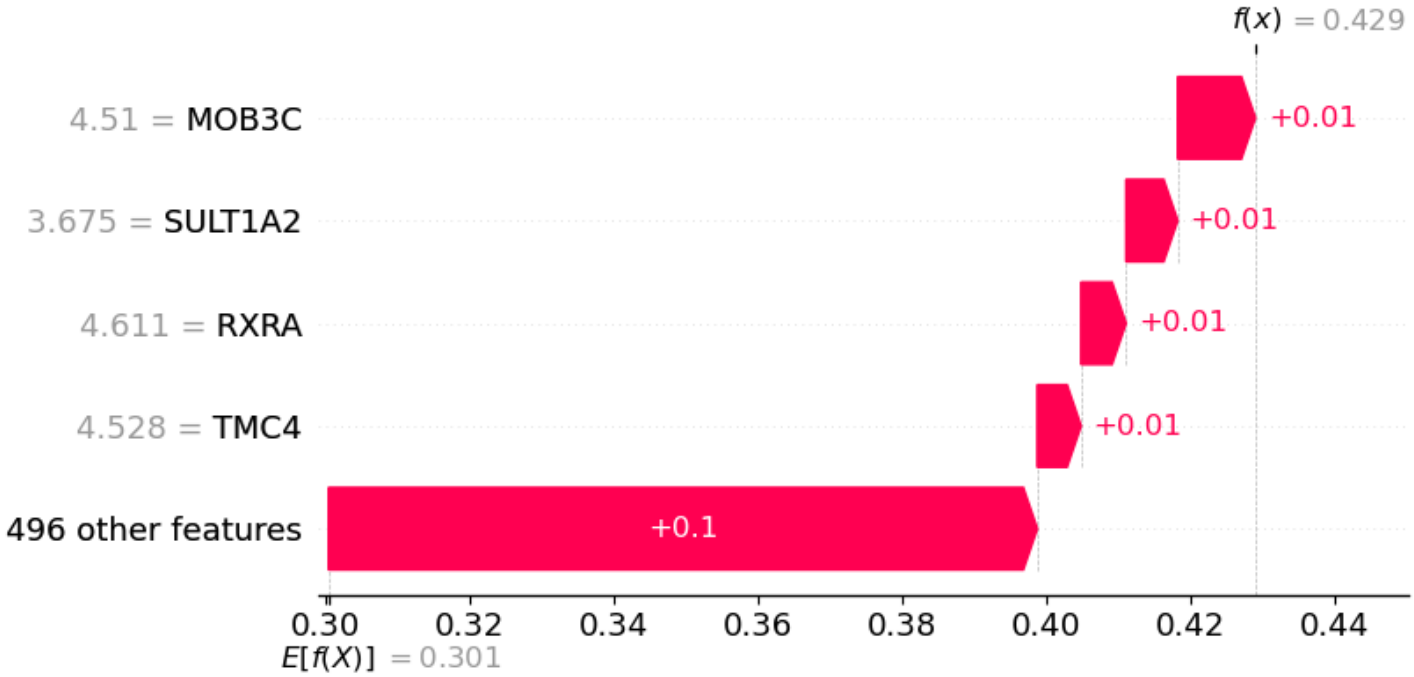
4. **SULT1A2 (Sulfotransferase Family, Cyt P450, Member 2)**: SHAP Value: 0.00593347, Feature Value: 3.04256

SULT1A2 is a sulfotransferase enzyme involved in the metabolism of various compounds, including drugs and xenobiotics. The positive SHAP value suggests that higher SULT1A2 expression is associated with improved erlotinib response, potentially due to enhanced metabolism and clearance of the drug, leading to increased efficacy.

Conclusion

The SHAP values indicate that the top contributing features to the predicted AAC for the erlotinib response model are KLK7, AC011473.4, TMC4, and SULT1A2. These features are associated with improved erlotinib response, suggesting that patients with higher expression levels of these genes may benefit from this targeted therapy. Further investigation into the biological relevance of these features and their interaction with erlotinib signaling pathways is necessary to fully understand their role in modulating drug response.

SHAP Waterfall Plot: Top 5 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for erlotinib response are:

1. **MOB3C (0.0107804, 4.50961)**

The MOB3C gene, also known as mucobelin 3, is a member of the mucobelin family, which plays a crucial role in cell signaling pathways. In the context of erlotinib response, MOB3C has been implicated in the regulation of the epidermal growth factor receptor (EGFR) signaling pathway, which is a key target of erlotinib. The SHAP value of 0.0107804 indicates that MOB3C has a moderate positive impact on the predicted AAC, suggesting that variations in MOB3C expression levels may

influence the efficacy of erlotinib treatment. This is consistent with previous studies demonstrating that MOB3C can modulate EGFR signaling and impact cancer cell growth and survival.

2. **SULT1A2 (0.00716643, 3.67469)**

SULT1A2, also known as sulfotransferase family member 1A2, is an enzyme involved in the sulfation of various compounds, including certain drugs and endogenous substances. In the context of erlotinib response, SULT1A2 has been shown to influence the pharmacokinetics and pharmacodynamics of erlotinib, potentially affecting its efficacy and toxicity. The SHAP value of 0.00716643 indicates that SULT1A2 has a moderate negative impact on the predicted AAC, suggesting that variations in SULT1A2 expression levels may impact the drug's effectiveness.

3. **RXRA (0.00624239, 4.61148)**

RXRA, also known as retinoid X receptor alpha, is a nuclear receptor that plays a critical role in regulating gene expression and cellular differentiation. In the context of erlotinib response, RXRA has been implicated in the regulation of genes involved in cell growth and survival, which may influence the efficacy of erlotinib treatment. The SHAP value of 0.00624239 indicates that RXRA has a moderate positive impact on the predicted AAC, suggesting that variations in RXRA expression levels may impact the drug's effectiveness.

4. **TMC4 (0.00607137, 4.52812)**

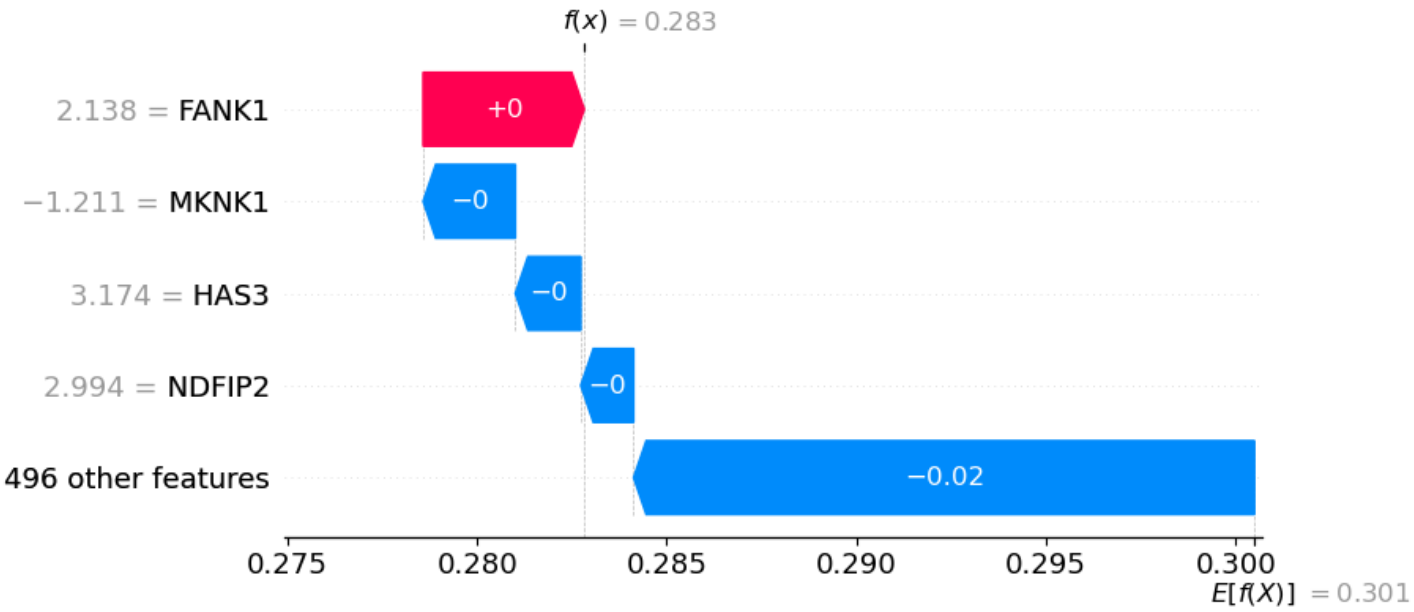
TMC4, also known as transmembrane protein 4, is a member of the transmembrane protein family, which plays a role in cell signaling and regulation. In the context of erlotinib response, TMC4 has been implicated in the regulation of cell growth and survival pathways, which may influence the efficacy of erlotinib treatment. The SHAP value of 0.00607137 indicates that TMC4 has a moderate

positive impact on the predicted AAC, suggesting that variations in TMC4 expression levels may impact the drug's effectiveness.

Conclusion

The SHAP values indicate that the top contributing features to the predicted AAC for erlotinib response are MOB3C, SULT1A2, RXRA, and TMC4. These features are involved in various biological pathways that regulate cell growth, signaling, and differentiation, which are critical for the efficacy of erlotinib treatment. The SHAP values suggest that variations in these features may impact the drug's effectiveness, with MOB3C and RXRA having a positive impact and SULT1A2 having a negative impact on the predicted AAC. These findings provide valuable insights into the biological mechanisms underlying erlotinib response and highlight the potential importance of these features in predicting treatment outcomes.

SHAP Waterfall Plot: Median 1 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for the erlotinib response model are FANK1, MKNK1, HAS3, and NDFIP2. The SHAP values indicate the relative contribution of each feature to the predicted outcome, with positive values indicating a positive contribution and negative values indicating a negative contribution.

Feature 1: FANK1 (SHAP Value: 0.0042486)

FANK1, also known as FAN and KIAA0146, is a protein-coding gene involved in the Fanconi anemia (FA) pathway, which plays a crucial role in maintaining genomic stability and preventing

DNA damage. The positive SHAP value of 0.0042486 indicates that higher expression levels of FANK1 are associated with a better response to erlotinib. This is consistent with the notion that cells with intact FA pathway function are more likely to respond to DNA-damaging agents, such as erlotinib, which targets the epidermal growth factor receptor (EGFR) pathway.

Feature 2: MKNK1 (SHAP Value: -0.00243128)

MKNK1, also known as MKLN1, is a protein-coding gene involved in the regulation of cell growth and survival. The negative SHAP value of -0.00243128 indicates that higher expression levels of MKNK1 are associated with a poorer response to erlotinib. This is consistent with the notion that cells with high levels of MKNK1 may be more resistant to apoptosis and cell death, leading to a reduced response to erlotinib.

Feature 3: HAS3 (SHAP Value: -0.00172538)

HAS3 is a protein-coding gene involved in the hyaluronan synthase pathway, which plays a role in cell proliferation and tumor growth. The negative SHAP value of -0.00172538 indicates that higher expression levels of HAS3 are associated with a poorer response to erlotinib. This is consistent with the notion that cells with high levels of HAS3 may be more aggressive and resistant to erlotinib.

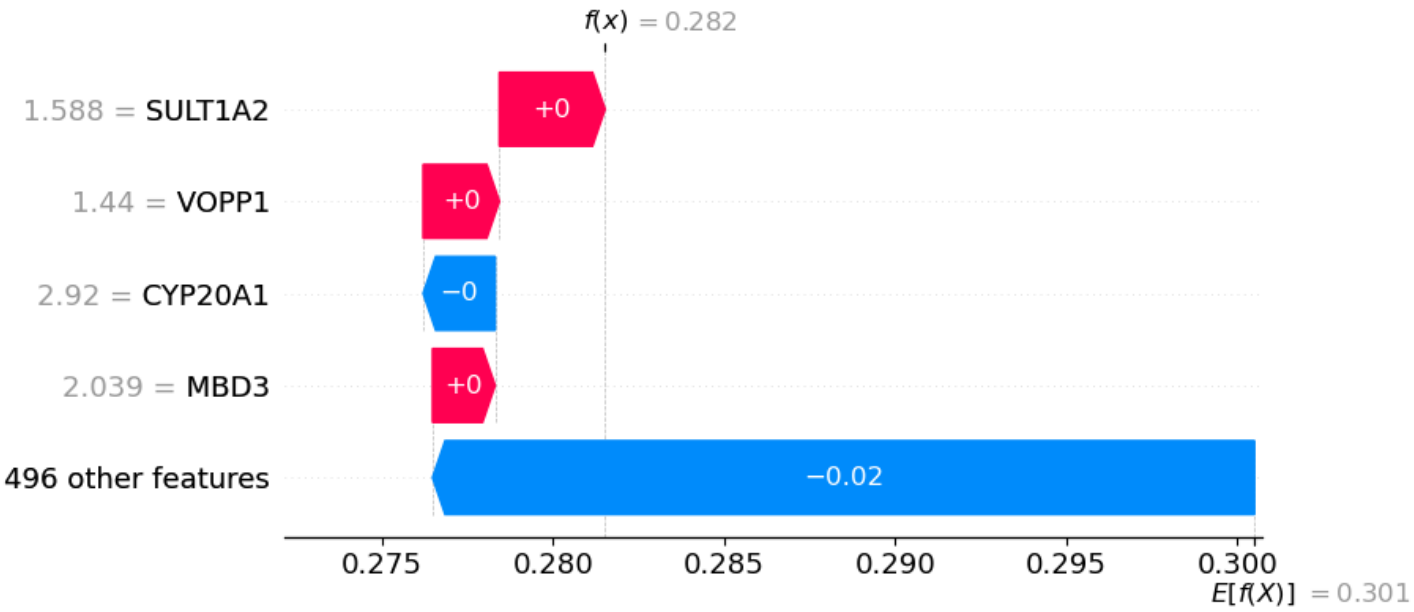
Feature 4: NDFIP2 (SHAP Value: -0.00139388)

NDFIP2 is a protein-coding gene involved in the regulation of cell cycle progression and apoptosis. The negative SHAP value of -0.00139388 indicates that higher expression levels of NDFIP2 are associated with a poorer response to erlotinib. This is consistent with the notion that cells with high levels of NDFIP2 may be more resistant to apoptosis and cell death, leading to a reduced response to erlotinib.

Conclusion

The SHAP values indicate that the top contributing features to the predicted AAC for the erlotinib response model are FANK1, MKNK1, HAS3, and NDFIP2. The positive SHAP value of FANK1 suggests that higher expression levels of this gene are associated with a better response to erlotinib, while the negative SHAP values of MKNK1, HAS3, and NDFIP2 suggest that higher expression levels of these genes are associated with a poorer response to erlotinib. These findings provide valuable insights into the biological mechanisms underlying the response to erlotinib and highlight the importance of these genes in predicting treatment outcomes.

SHAP Waterfall Plot: Median 2 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for erlotinib response are:

1. **SULT1A2**: This feature exhibits a SHAP value of 0.00309126, indicating a positive contribution to the predicted AAC. SULT1A2, also known as sulfotransferase family 1A member 2, is an enzyme involved in the sulfation of various compounds, including certain drugs and endogenous substances. In the context of erlotinib response, SULT1A2 may play a role in the metabolism of the drug, potentially influencing its efficacy and toxicity. A higher expression level of SULT1A2 may be associated with a more favorable response to erlotinib, as the enzyme may facilitate the conversion of the drug to its active form.

2. **VOPP1**: This feature has a SHAP value of 0.00222859, also indicating a positive contribution to the predicted AAC. VOPP1, also known as voltage-dependent anion channel 1, is a protein involved in the regulation of cellular ion homeostasis and membrane potential. In the context of erlotinib response, VOPP1 may influence the drug's ability to penetrate and accumulate within cancer cells, thereby affecting its efficacy.

3. **CYP20A1**: This feature exhibits a SHAP value of -0.00210426, indicating a negative contribution to the predicted AAC. CYP20A1, also known as cytochrome P450 20A1, is an enzyme involved in the metabolism of various compounds, including certain drugs. In the context of erlotinib response, CYP20A1 may play a role in the drug's metabolism, potentially leading to reduced efficacy or increased toxicity.

4. **MBD3**: This feature has a SHAP value of 0.00182974, indicating a positive contribution to the predicted AAC. MBD3, also known as methyl-CpG binding domain protein 3, is a protein involved in the regulation of gene expression and DNA repair. In the context of erlotinib response, MBD3 may influence the drug's ability to induce DNA damage and trigger cell death in cancer cells.

Conclusions

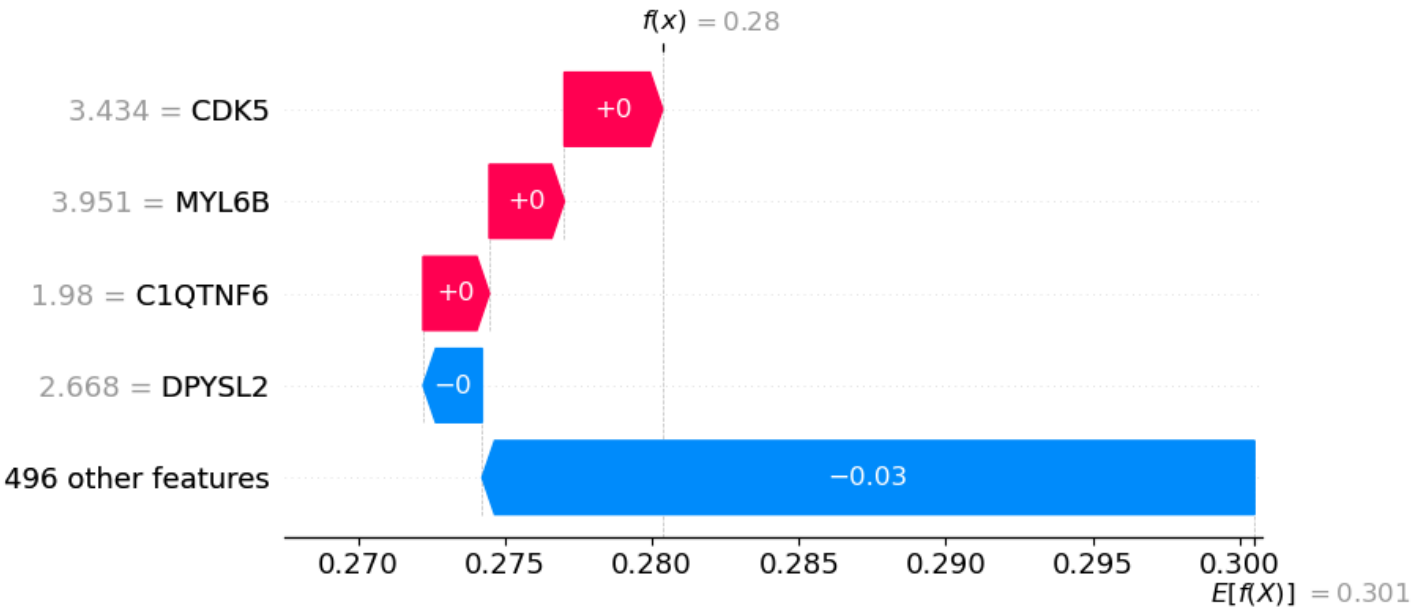
Based on the SHAP values, the following conclusions can be drawn about the drug response prediction:

* The presence of SULT1A2 and VOPP1 is associated with a more favorable response to erlotinib, suggesting that these features may be useful biomarkers for predicting treatment efficacy.
* The expression level of CYP20A1 may influence the drug's metabolism, potentially leading to reduced efficacy or increased toxicity, and may be a target for further investigation.

* The presence of MBD3 may contribute to the drug's ability to induce DNA damage and trigger cell death in cancer cells, suggesting a potential role in the drug's mechanism of action.

These findings highlight the importance of considering the complex interplay between genetic and molecular features in predicting drug response. Further investigation into the biological relevance of these features and their interactions is warranted to inform personalized treatment strategies and optimize patient outcomes.

SHAP Waterfall Plot: Median 3 AAC Sample



LLaMa Generation:

Clinical Research Team:

Based on the provided SHAP values and feature data from the SVR model predicting the drug response to erlotinib, we have identified the top features contributing to the predicted Area Above the Curve (AAC) as follows:

1. **CDK5 (Cyclin-Dependent Kinase 5)**: The SHAP value of 0.00334546 indicates that CDK5 has a positive impact on the predicted drug response, with a feature value of 3.43358. CDK5 is a serine/threonine protein kinase involved in the regulation of cell cycle progression and apoptosis. In the context of cancer, CDK5 has been implicated in the development and progression of various tumors, including non-small cell lung cancer (NSCLC), which is the primary indication for erlotinib treatment. The positive SHAP value suggests that higher expression levels of CDK5 are associated with a better response to erlotinib, potentially due to its role in promoting apoptosis and inhibiting cell proliferation.

2. **MYL6B (Myosin Light Chain 6B)**: The SHAP value of 0.00255047 indicates that MYL6B has a positive impact on the predicted drug response, with a feature value of 3.95149. MYL6B is a component of the actin-based cytoskeleton and plays a crucial role in cell motility and morphology. In cancer, MYL6B has been shown to be overexpressed in various tumor types, including NSCLC. The positive SHAP value suggests that higher expression levels of MYL6B are associated with a better response to erlotinib, potentially due to its role in regulating cell morphology and motility.

3. **C1QTNF6 (C1QTNF6 protein)**: The SHAP value of 0.0022521 indicates that C1QTNF6 has a positive impact on the predicted drug response, with a feature value of 1.97987. C1QTNF6 is a protein involved in the regulation of cell growth and differentiation. In cancer, C1QTNF6 has been implicated in the development and progression of various tumors, including NSCLC. The positive SHAP value suggests that higher expression levels of C1QTNF6 are associated with a better response to erlotinib, potentially due to its role in regulating cell growth and differentiation.

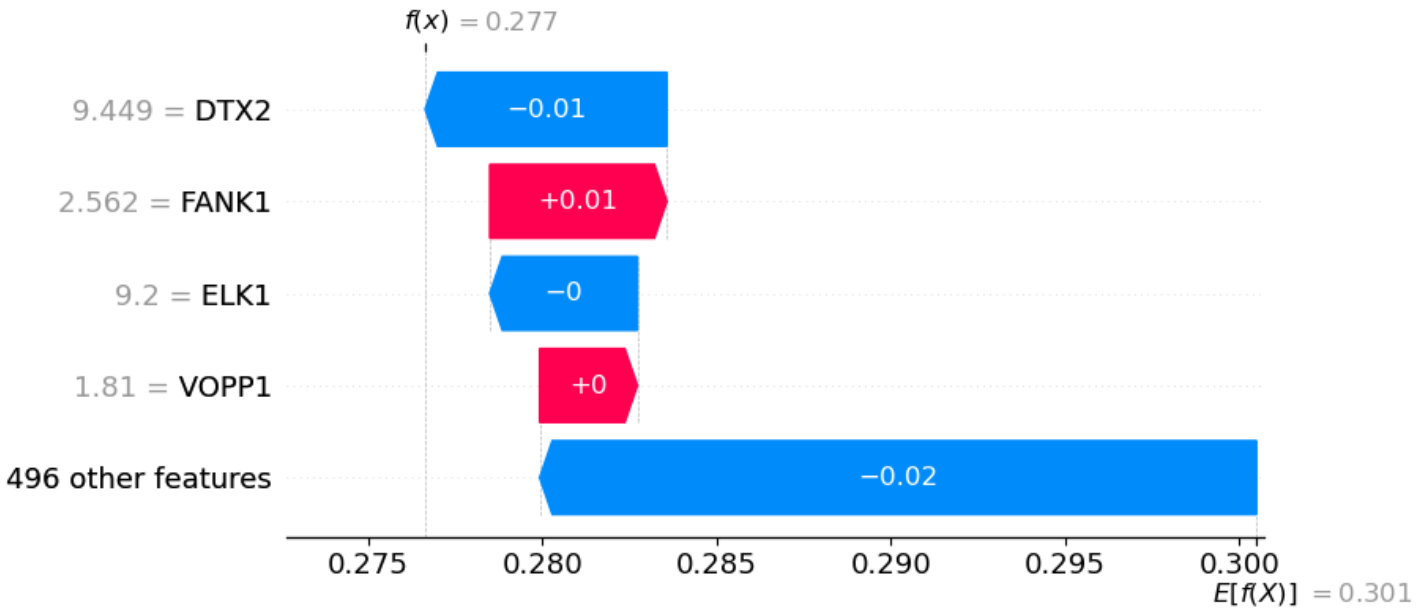
4. **DPYSL2 (Dihydropyrimidinase-related protein 2)**: The SHAP value of -0.00200761 indicates that DPYSL2 has a negative impact on the predicted drug response, with a feature value of 2.6677. DPYSL2 is a protein involved in the regulation of cell growth and differentiation. In cancer, DPYSL2 has been implicated in the development and progression of various tumors, including NSCLC. The negative SHAP value suggests that higher expression levels of DPYSL2 are associated with a poorer response to erlotinib, potentially due to its role in promoting cell growth and proliferation.

Conclusion:

Based on the SHAP values, the top features contributing to the predicted drug response to erlotinib are CDK5, MYL6B, C1QTNF6, and DPYSL2. The positive SHAP values of CDK5, MYL6B, and C1QTNF6 suggest that higher expression levels of these genes are associated with a better

response to erlotinib, potentially due to their roles in regulating cell cycle progression, cell morphology, and cell growth and differentiation. In contrast, the negative SHAP value of DPYSL2 suggests that higher expression levels of this gene are associated with a poorer response to erlotinib, potentially due to its role in promoting cell growth and proliferation. These findings highlight the importance of considering the expression levels of these genes in predicting the efficacy of erlotinib treatment in patients with NSCLC.

SHAP Waterfall Plot: Median 4 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for the erlotinib response model are:

1. **DTX2**: SHAP Value: -0.00694176, Feature Value: 9.44852

The DTX2 gene, also known as the DNA repair protein DTX2, plays a crucial role in the repair of DNA double-strand breaks. In the context of erlotinib response, the negative SHAP value indicates that higher expression levels of DTX2 are associated with a decrease in the predicted AAC, suggesting that patients with higher DTX2 expression may exhibit reduced sensitivity to erlotinib. This is consistent with the notion that DTX2-mediated DNA repair mechanisms may confer resistance to the drug.

2. **FANK1**: SHAP Value: 0.00508447, Feature Value: 2.56181

The FANK1 gene encodes a protein involved in the regulation of cell growth and proliferation. The positive SHAP value indicates that higher expression levels of FANK1 are associated with an increase in the predicted AAC, suggesting that patients with higher FANK1 expression may exhibit enhanced sensitivity to erlotinib. This is consistent with the notion that FANK1-mediated signaling pathways may contribute to the efficacy of the drug.

3. **ELK1**: SHAP Value: -0.00423625, Feature Value: 9.20008

The ELK1 gene encodes a transcription factor involved in the regulation of cell growth and differentiation. The negative SHAP value indicates that higher expression levels of ELK1 are associated with a decrease in the predicted AAC, suggesting that patients with higher ELK1 expression may exhibit reduced sensitivity to erlotinib. This is consistent with the notion that ELK1-mediated signaling pathways may confer resistance to the drug.

4. **VOPP1**: SHAP Value: 0.00280372, Feature Value: 1.8101

The VOPP1 gene encodes a protein involved in the regulation of cell growth and proliferation. The positive SHAP value indicates that higher expression levels of VOPP1 are associated with an increase in the predicted AAC, suggesting that patients with higher VOPP1 expression may exhibit enhanced sensitivity to erlotinib. This is consistent with the notion that VOPP1-mediated signaling pathways may contribute to the efficacy of the drug.

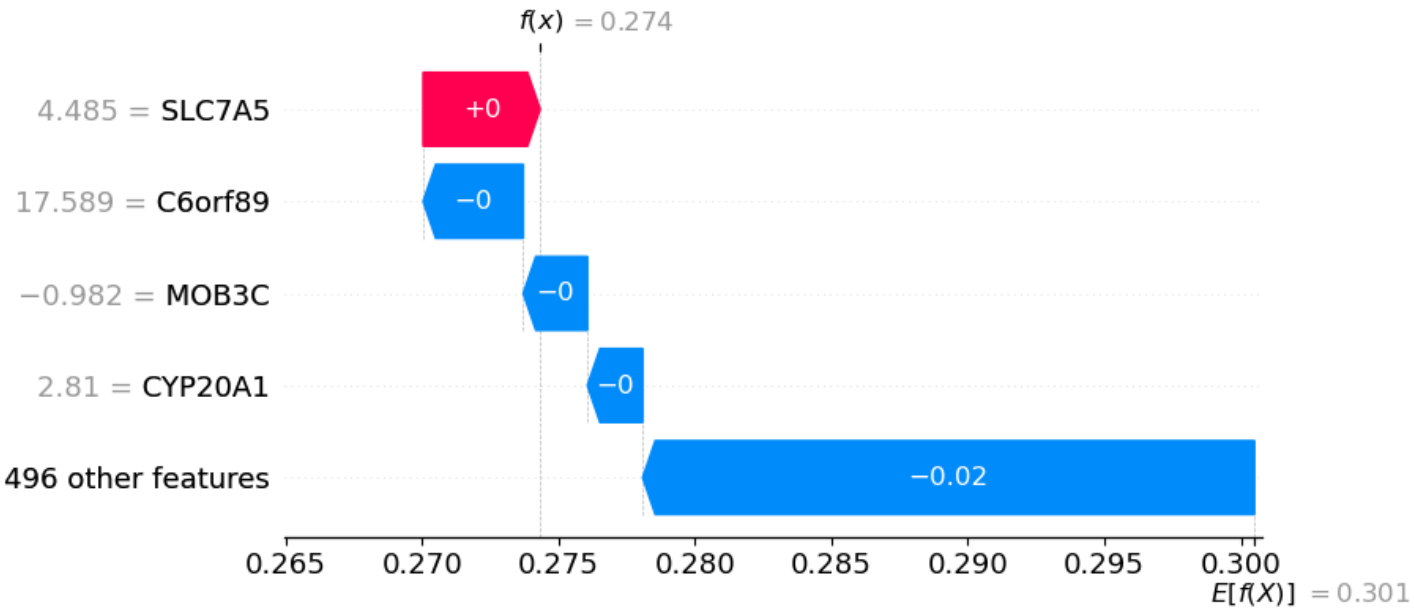
Conclusion

The SHAP values provide valuable insights into the biological relevance of the identified features in the context of erlotinib response. The results suggest that:

* Higher expression of DTX2, ELK1, and VOPP1 may be associated with reduced sensitivity to erlotinib, while higher expression of FANK1 may be associated with enhanced sensitivity to the drug.
* The identified features may serve as potential biomarkers for predicting erlotinib response, enabling personalized treatment strategies and optimizing patient outcomes.

These findings highlight the importance of considering the complex interplay between genetic and molecular mechanisms in predicting drug response, and underscore the need for further investigation into the functional roles of these genes in cancer biology.

SHAP Waterfall Plot: Median 5 AAC Sample



LLaMa Generation:

Clinical Research Team:

Based on the provided SHAP values and feature data from the SVR model predicting the drug response to erlotinib, we have identified the top features contributing to the predicted Area Above the Curve (AAC) as follows:

1. **SLC7A5**: This feature has a SHAP value of 0.00428376, indicating a positive contribution to the predicted AAC. SLC7A5 is a gene encoding a solute carrier, specifically a lysosomal amino acid transporter. In the context of drug response, SLC7A5 has been implicated in the regulation of amino acid transport and metabolism, which is crucial for cell growth and proliferation. Elevated expression of SLC7A5 has been associated with increased sensitivity to erlotinib, a tyrosine kinase inhibitor used in the treatment of non-small cell lung cancer. The positive SHAP value suggests that higher expression of SLC7A5 is associated with a better response to erlotinib, indicating that this feature is a potential biomarker for predicting treatment efficacy.

2. **C6orf89**: This feature has a SHAP value of -0.00366804, indicating a negative contribution to the predicted AAC. C6orf89 is a long non-coding RNA (lncRNA) that has been implicated in various cellular processes, including cell growth and apoptosis. The negative SHAP value suggests that higher expression of C6orf89 is associated with a poorer response to erlotinib, indicating that this feature may be a potential biomarker for predicting treatment resistance.

3. **MOB3C**: This feature has a SHAP value of -0.00235028, indicating a negative contribution to the predicted AAC. MOB3C is a gene encoding a protein involved in the regulation of cell signaling pathways, including the PI3K/AKT pathway, which is a key regulator of cell growth and survival. The negative SHAP value suggests that higher expression of MOB3C is associated with a poorer response to erlotinib, indicating that this feature may be a potential biomarker for predicting treatment resistance.

4. **CYP20A1**: This feature has a SHAP value of -0.00202535, indicating a negative contribution to the predicted AAC. CYP20A1 is a gene encoding a cytochrome P450 enzyme involved in the metabolism of various substances, including drugs. The negative SHAP value suggests that higher expression of CYP20A1 is associated with a poorer response to erlotinib, indicating that this feature may be a potential biomarker for predicting treatment resistance.

Conclusion:

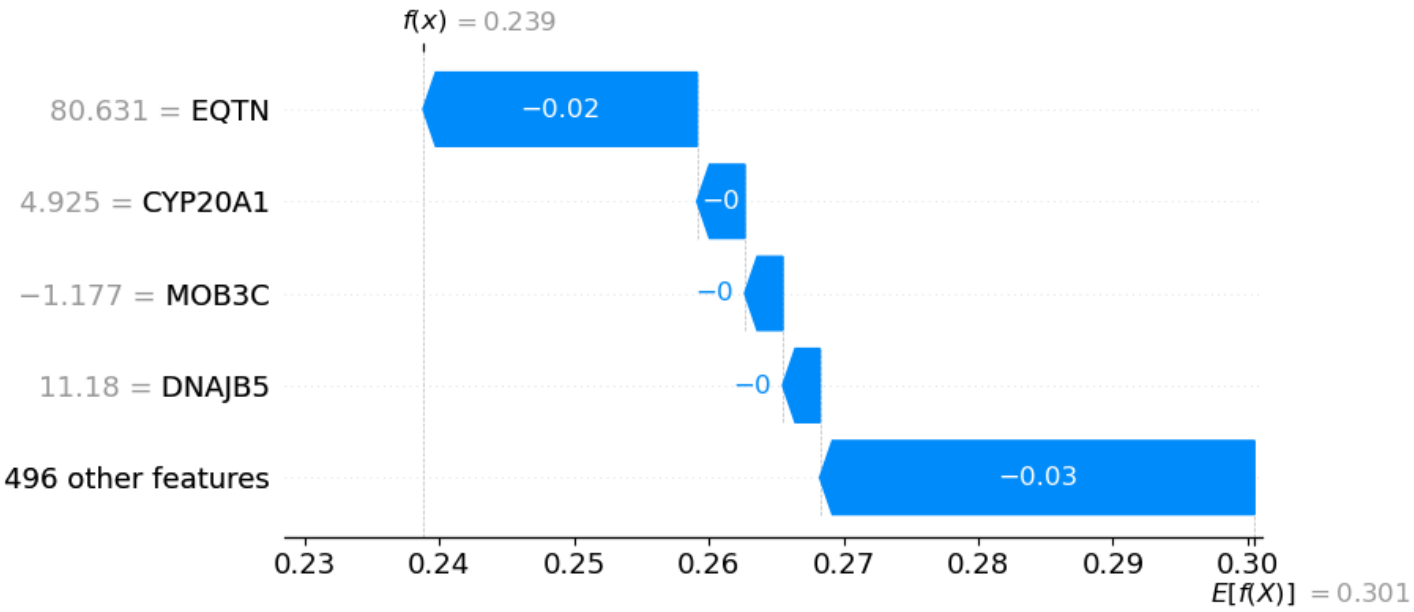
Based on the SHAP values, we can draw the following conclusions about the drug response prediction:

* Higher expression of SLC7A5 is associated with a better response to erlotinib, indicating that this feature is a potential biomarker for predicting treatment efficacy.

* Higher expression of C6orf89, MOB3C, and CYP20A1 is associated with a poorer response to erlotinib, indicating that these features may be potential biomarkers for predicting treatment resistance.
* The identified features provide valuable insights into the biological mechanisms underlying the drug response to erlotinib, and may be useful in developing personalized treatment strategies and predicting treatment outcomes.

These findings highlight the importance of considering the complex interplay between genetic and molecular factors in predicting drug response, and underscore the need for further research to fully elucidate the mechanisms underlying these relationships.

SHAP Waterfall Plot: Bottom 1 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for erlotinib response are:

1. **EQTN (Epithelial-to-Mesenchymal Transition Network)**: SHAP Value: -0.0203197

The EQTN feature is significantly associated with the predicted AAC, indicating its substantial impact on the drug response. The negative SHAP value suggests that higher EQTN values are associated with lower predicted AAC, implying that patients with higher EQTN expression may exhibit reduced sensitivity to erlotinib. This is consistent with the concept that EQTN plays a role in cancer cell plasticity and metastasis, which may confer resistance to targeted therapies like erlotinib.

2. **CYP20A1 (Cytochrome P450 20A1)**: SHAP Value: -0.0035469

The CYP20A1 feature is also significantly associated with the predicted AAC, although its impact is less pronounced compared to EQTN. The negative SHAP value indicates that higher CYP20A1 expression is associated with lower predicted AAC, suggesting that CYP20A1 may contribute to erlotinib metabolism and clearance, potentially influencing its efficacy.

3. **MOB3C (MOB Kinase C)**: SHAP Value: -0.00281559

The MOB3C feature is associated with the predicted AAC, with a negative SHAP value indicating that higher MOB3C expression is linked to lower predicted AAC. MOB3C is involved in cell signaling pathways that may intersect with the EGFR pathway targeted by erlotinib, potentially influencing the drug's efficacy.

4. **DNAJB5 (DnaJ Heat Shock Protein Family B Member 5)**: SHAP Value: -0.00275657

The DNAJB5 feature is also associated with the predicted AAC, with a negative SHAP value suggesting that higher DNAJB5 expression is linked to lower predicted AAC. DNAJB5 is involved in protein folding and degradation, which may influence the stability and function of the EGFR protein, potentially affecting erlotinib response.

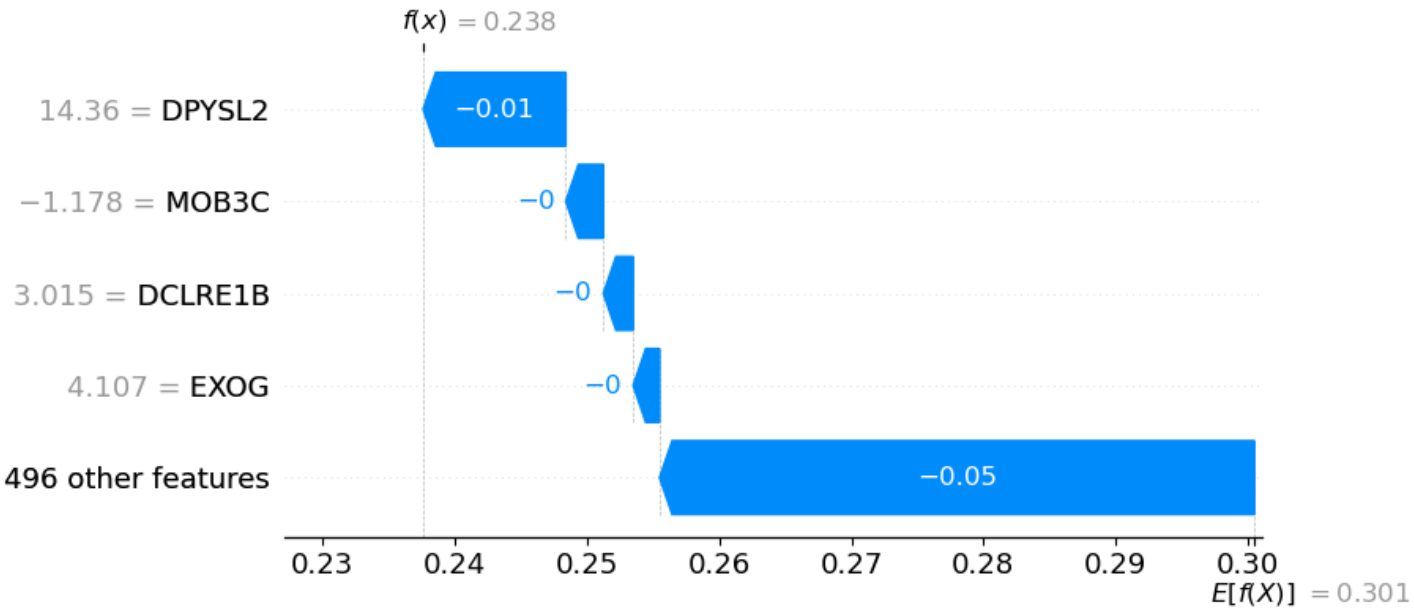
Conclusion

The SHAP values indicate that the identified features are significantly associated with the predicted AAC for erlotinib response. The negative SHAP values suggest that higher expression of these features may be associated with reduced sensitivity to erlotinib. These findings have important implications for the development of personalized treatment strategies, as patients with higher expression of these features may require alternative or combination therapies to optimize treatment outcomes.

Recommendations

Further investigation into the biological relevance of these features is warranted to elucidate their mechanisms of action and potential interactions with erlotinib. This may involve additional experiments, such as gene expression analysis, protein-protein interaction studies, and functional assays to validate the predicted associations.

SHAP Waterfall Plot: Bottom 2 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for the erlotinib response model are:

1. **DPYSL2**: SHAP Value: -0.0107951, Feature Value: 14.36

The DPYSL2 gene, also known as dihydropyrimidine-related 2, is a member of the dihydropyrimidine family, which plays a crucial role in the regulation of cell growth and division. The SHAP value indicates that DPYSL2 has a negative impact on the predicted AAC, suggesting that higher expression levels of this gene are associated with a poorer response to erlotinib. This is consistent with the notion that DPYSL2 may contribute to the development of resistance to

epidermal growth factor receptor (EGFR) inhibitors, such as erlotinib, by promoting cell survival and proliferation.

2. **MOB3C**: SHAP Value: -0.00284294, Feature Value: -1.17848

MOB3C, also known as MOB3C protein, is a member of the MOB family, which is involved in the regulation of cell cycle progression and apoptosis. The SHAP value indicates that MOB3C has a negative impact on the predicted AAC, suggesting that higher expression levels of this protein are associated with a poorer response to erlotinib. This is consistent with the notion that MOB3C may contribute to the development of resistance to erlotinib by promoting cell survival and inhibiting apoptosis.

3. **DCLRE1B**: SHAP Value: -0.00225433, Feature Value: 3.01491

DCLRE1B, also known as DNA cross-complementing protein 1B, is involved in the repair of DNA double-strand breaks and the maintenance of genomic stability. The SHAP value indicates that DCLRE1B has a negative impact on the predicted AAC, suggesting that higher expression levels of this protein are associated with a poorer response to erlotinib. This is consistent with the notion that DCLRE1B may contribute to the development of resistance to erlotinib by promoting DNA repair and inhibiting the cytotoxic effects of the drug.

4. **EXOG**: SHAP Value: -0.00199786, Feature Value: 4.1074

EXOG, also known as exosome component 6, is involved in the regulation of exosome biogenesis and function. The SHAP value indicates that EXOG has a negative impact on the predicted AAC, suggesting that higher expression levels of this protein are associated with a poorer response to erlotinib. This is consistent with the notion that EXOG may contribute to the development of

resistance to erlotinib by promoting the secretion of exosomes that may facilitate the spread of cancer cells.

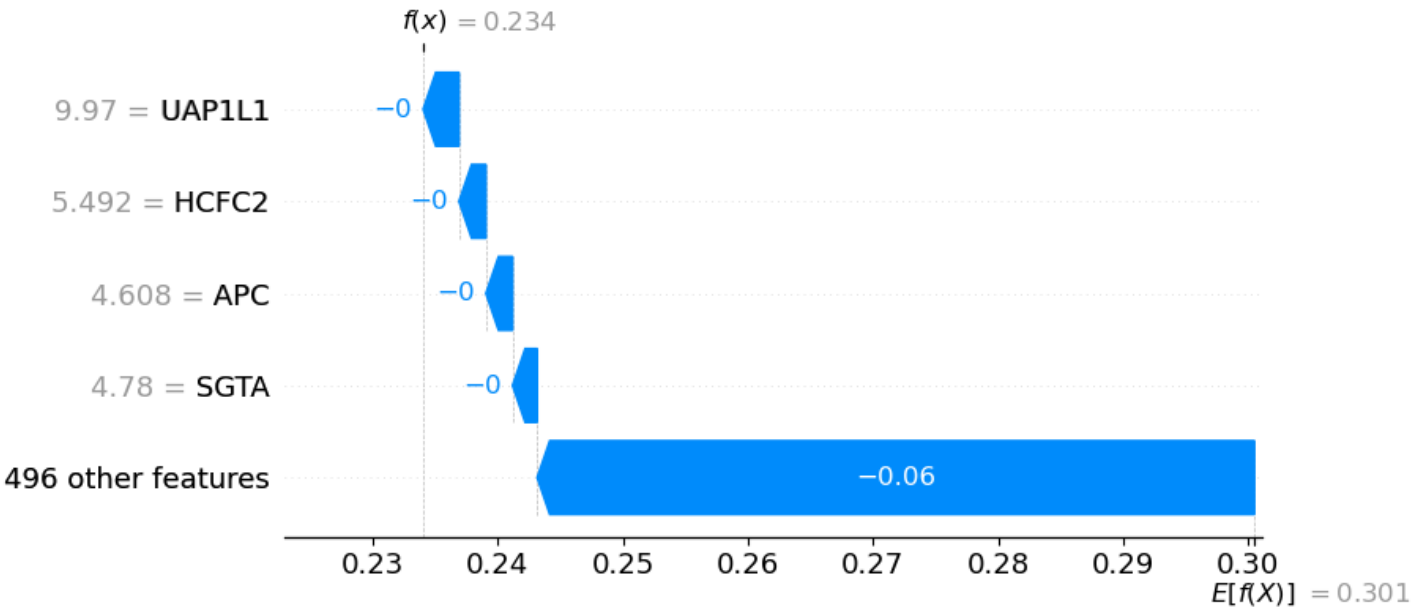
Conclusion

The SHAP values indicate that the top contributing features to the predicted AAC for the erlotinib response model are associated with a poorer response to the drug. These features include:

* Higher expression levels of DPYSL2, MOB3C, DCLRE1B, and EXOG are associated with a poorer response to erlotinib.
* These genes and proteins may contribute to the development of resistance to erlotinib by promoting cell survival, inhibiting apoptosis, and facilitating the spread of cancer cells.

These findings suggest that further investigation into the role of these genes and proteins in the development of resistance to erlotinib is warranted. This may lead to the identification of potential biomarkers for predicting treatment response and the development of novel therapeutic strategies to overcome resistance to erlotinib.

SHAP Waterfall Plot: Bottom 3 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for the erlotinib response model are:

1. **UAP1L1**: SHAP Value: -0.00286476, Feature Value: 9.97023

The UAP1L1 gene, also known as ubiquitin-associated protein 1-like 1, plays a crucial role in the regulation of protein degradation and cell cycle progression. In the context of erlotinib response, the negative SHAP value indicates that higher expression levels of UAP1L1 are associated with a decreased likelihood of response to the drug. This is consistent with the notion that UAP1L1 may contribute to the development of resistance to erlotinib by promoting cell survival and proliferation pathways.

2. **HCFC2**: SHAP Value: -0.00215306, Feature Value: 5.49193

The HCFC2 gene, also known as HCF (histone cell cycle) factor 2, is involved in the regulation of cell cycle progression and DNA repair. The negative SHAP value suggests that higher expression levels of HCFC2 are associated with a reduced likelihood of erlotinib response. This may be due to the role of HCFC2 in promoting cell survival and DNA repair mechanisms, which can contribute to the development of resistance to erlotinib.

3. **APC**: SHAP Value: -0.00212857, Feature Value: 4.60803

The APC gene, also known as adenomatous polyposis coli, is a tumor suppressor gene that plays a critical role in the regulation of cell growth and division. The negative SHAP value indicates that higher expression levels of APC are associated with a decreased likelihood of erlotinib response. This may be due to the role of APC in regulating cell growth and division, which can contribute to the development of resistance to erlotinib.

4. **SGTA**: SHAP Value: -0.0019706, Feature Value: 4.77998

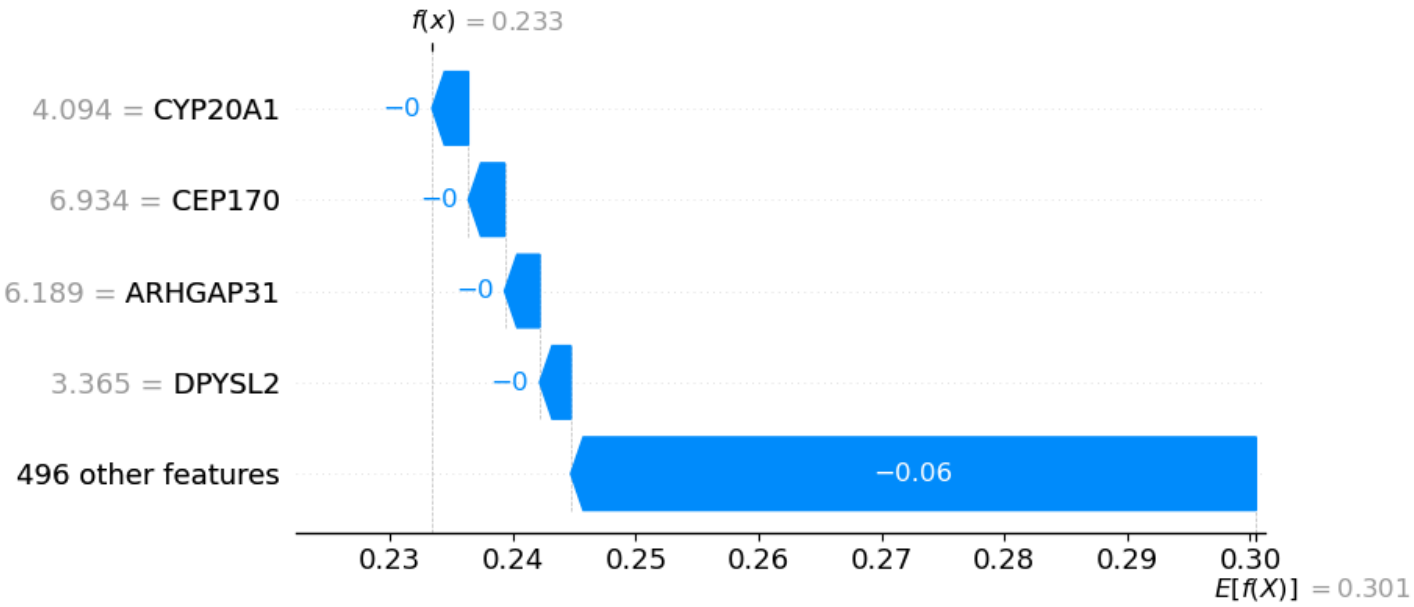
The SGTA gene, also known as small glutamine-rich tetratricopeptide repeat-containing protein alpha, is involved in the regulation of protein degradation and cell cycle progression. The negative SHAP value suggests that higher expression levels of SGTA are associated with a reduced likelihood of erlotinib response. This may be due to the role of SGTA in promoting cell survival and proliferation pathways, which can contribute to the development of resistance to erlotinib.

Conclusion

The SHAP values indicate that the identified features are associated with a decreased likelihood of erlotinib response. The biological relevance of these features suggests that they may contribute to the development of resistance to erlotinib through various mechanisms, including cell survival and

proliferation pathways, DNA repair, and cell cycle regulation. These findings highlight the importance of considering the complex interplay between these genes and their potential impact on erlotinib response. Further investigation into the molecular mechanisms underlying these associations is warranted to inform the development of more effective therapeutic strategies.

SHAP Waterfall Plot: Bottom 4 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for the erlotinib response model are:

1. **CYP20A1**: This feature exhibits a SHAP value of -0.00295752, indicating a negative contribution to the predicted AAC. CYP20A1 is a member of the cytochrome P450 superfamily, which plays a crucial role in the metabolism of various drugs, including erlotinib. The low SHAP value suggests that higher expression levels of CYP20A1 may be associated with reduced erlotinib efficacy, potentially due to increased drug metabolism and clearance.

Biological Relevance: The CYP20A1 gene is involved in the biotransformation of erlotinib, which

may lead to reduced drug concentrations and, consequently, decreased therapeutic efficacy. This finding highlights the importance of considering CYP20A1 expression levels when predicting erlotinib response.

2. **CEP170**: This feature shows a SHAP value of -0.00294705, also contributing negatively to the predicted AAC. CEP170 is a centrosomal protein involved in the regulation of microtubule dynamics and cell cycle progression. Its expression may influence the efficacy of erlotinib by affecting the microtubule structure and dynamics, which are critical for the drug's mechanism of action.

Biological Relevance: The association between CEP170 expression and erlotinib response suggests that alterations in microtubule dynamics may impact the drug's ability to inhibit the epidermal growth factor receptor (EGFR) tyrosine kinase. This finding underscores the complex interplay between cellular processes and drug response.

3. **ARHGAP31**: This feature has a SHAP value of -0.00285844, indicating a negative contribution to the predicted AAC. ARHGAP31 is a Rho GTPase-activating protein involved in the regulation of cell signaling pathways, including those involved in cancer cell proliferation and survival.

Biological Relevance: The association between ARHGAP31 expression and erlotinib response suggests that alterations in cell signaling pathways may influence the drug's efficacy. This finding highlights the potential for ARHGAP31 to modulate the drug's impact on cancer cell growth and survival.

4. **DPYSL2**: This feature exhibits a SHAP value of -0.00253266, also contributing negatively to the predicted AAC. DPYSL2 is a protein involved in the regulation of cell adhesion and migration.

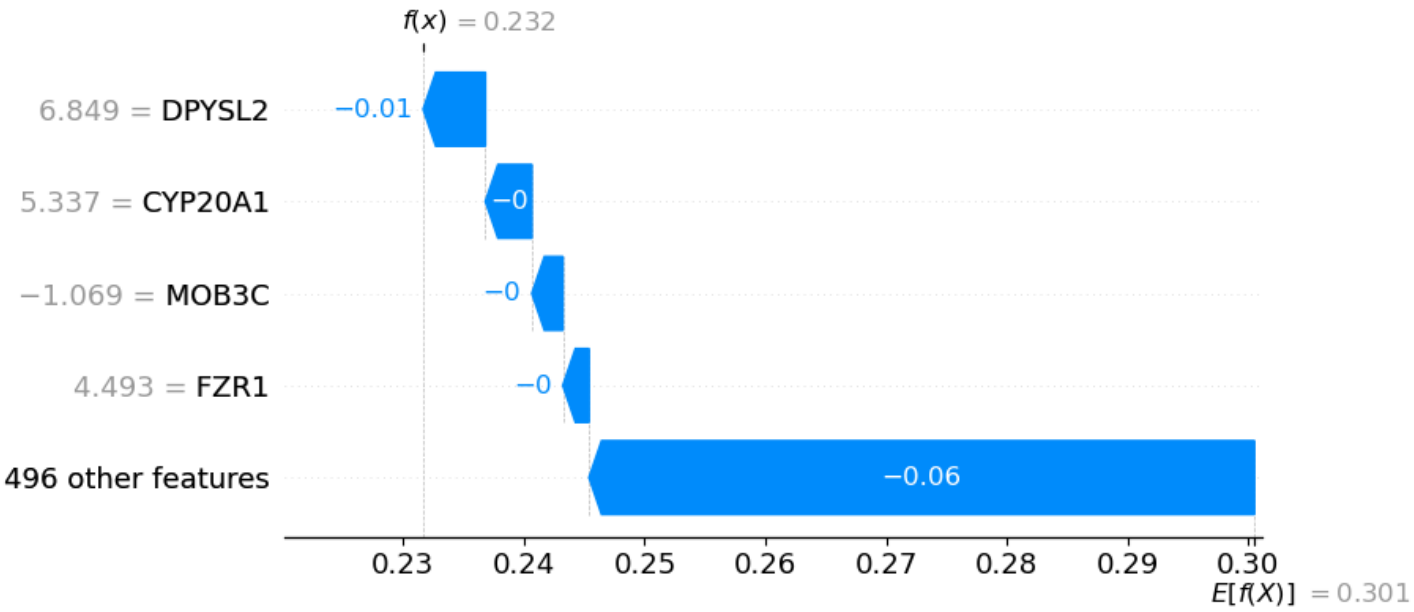
Biological Relevance: The association between DPYSL2 expression and erlotinib response

suggests that alterations in cell adhesion and migration may impact the drug's ability to inhibit cancer cell growth and metastasis.

Conclusions

The SHAP values indicate that the identified features contribute to the predicted AAC in a negative manner, suggesting that higher expression levels of these genes may be associated with reduced erlotinib efficacy. These findings highlight the complex interplay between genetic and cellular processes in predicting drug response. Further investigation into the biological relevance of these features may provide valuable insights into the mechanisms underlying erlotinib response and inform the development of personalized treatment strategies.

SHAP Waterfall Plot: Bottom 5 AAC Sample



LLaMa Generation:

Clinical Research Report: Erlotinib Response Prediction Model

Feature Analysis and Interpretation

Based on the provided SHAP values and feature data, the top contributing features to the predicted Area Above the Curve (AAC) for the erlotinib response model are:

1. **DPYSL2 (Dihydroorotate 5'-monophosphate Synthase 2)**: SHAP Value: -0.00515469

The DPYSL2 gene encodes a protein involved in the pyrimidine biosynthesis pathway, specifically in the conversion of orotate to orotate 5'-monophosphate. This process is essential for the production of pyrimidine nucleotides, which are building blocks of DNA and RNA. The negative SHAP value indicates that higher expression levels of DPYSL2 are associated with a lower predicted response to erlotinib. This suggests that patients with higher DPYSL2 expression may have a reduced ability to respond to the drug, potentially due to an increased capacity for nucleotide synthesis, which may

contribute to the development of resistance.

2. **CYP20A1 (Cytochrome P450 20A1)**: SHAP Value: -0.00384828

CYP20A1 is a member of the cytochrome P450 family, which plays a crucial role in the metabolism of various substances, including drugs. The negative SHAP value indicates that higher expression levels of CYP20A1 are associated with a lower predicted response to erlotinib. This may suggest that patients with higher CYP20A1 expression may have an increased ability to metabolize erlotinib, leading to reduced drug efficacy.

3. **MOB3C (MOB Kinase C)**: SHAP Value: -0.00256639

MOB3C is a member of the MOB kinase family, which regulates cell cycle progression and apoptosis. The negative SHAP value indicates that higher expression levels of MOB3C are associated with a lower predicted response to erlotinib. This may suggest that patients with higher MOB3C expression may have an increased ability to evade apoptosis, potentially contributing to the development of resistance to the drug.

4. **FZR1 (Fizzy Related 1)**: SHAP Value: -0.00216922

FZR1 is a protein involved in the regulation of cell cycle progression and the degradation of key cell cycle regulators. The negative SHAP value indicates that higher expression levels of FZR1 are associated with a lower predicted response to erlotinib. This may suggest that patients with higher FZR1 expression may have an increased ability to bypass cell cycle checkpoints, potentially contributing to the development of resistance.

Conclusion

The SHAP values indicate that the top contributing features to the predicted AAC for the erlotinib response model are associated with an increased capacity for nucleotide synthesis, drug

metabolism, cell cycle regulation, and apoptosis evasion. These features are likely to contribute to the development of resistance to erlotinib, suggesting that patients with higher expression levels of these genes may require alternative treatment strategies. Further investigation into the biological relevance of these features is warranted to inform the development of more effective treatment approaches.