DARE: Towards Robust Text Explanations in Biomedical and Healthcare Applications

Adam Ivankay
IBM Research Zurich
Rüschlikon, Switzerland
aiv@zurich.ibm.com

Mattia Rigotti
IBM Research Zurich
Rüschlikon, Switzerland
mrg@zurich.ibm.com

Pascal Frossard
École Polytechnique Fédérale de Lausanne (EPFL)
Lausanne, Switzerland
pascal.frossard@epfl.ch

Abstract

Along with the successful deployment of deep neural networks in several application domains, the need to unravel the black-box nature of these networks has seen a significant increase recently. Several methods have been introduced to provide insight into the inference process of deep neural networks. However, most of these explainability methods have been shown to be brittle in the face of adversarial perturbations of their inputs in the image and generic textual domain. In this work we show that this phenomenon extends to specific and important high stakes domains like biomedical datasets. In particular, we observe that the robustness of explanations should be characterized in terms of the accuracy of the explanation in linking a model's inputs and its decisions – *faithfulness* – and its relevance from the perspective of domain experts – *plausibility*. This is crucial to prevent explanations that are inaccurate but still look convincing in the context of the domain at hand. To this end, we show how to adapt current attribution robustness estimation methods to a given domain, so as to take into account domain-specific plausibility. This results in our DOMAINADAPTIVEARESTIMATOR (DARE) attribution robustness estimator, allowing us to properly characterize the domain-specific robustness of faithful explanations. Next, we provide two methods, adversarial training and FAR training, to mitigate the brittleness characterized by DARE, allowing us to train networks that display robust attributions. Finally, we empirically validate our methods with extensive experiments on three established biomedical benchmarks.

1 Introduction

Research in explainable AI (XAI) has seen a surge in recent years. XAI methods aim to provide insight into the inference process and the causal links between inputs and outputs of deep neural networks (DNNs). This is pivotal in addressing many aspects of DNNs, such as fairness, potential biases and scopes of safe deployment. Especially in safety-critical domains, such as healthcare, faithful and robust explanations (Jacovi and Goldberg, 2020) accompanying the predictions of DNNs are key to enable their deployment and understand potential false predictions and risks. Beside faithfulness, which quantifies the accuracy by which the explanations characterize the true decision-making process of the model, a second property of explanations that has been highlighted as important is plausibility (Jacovi and Goldberg, 2021; Rigotti et al., 2022; Rizzo et al., 2022). Plausibility quantifies the "degree to which some explanation is aligned with the user's understanding" (Jacovi and Goldberg, 2021). As such, plausibility tells us whether an explanation is found convincing and informative for domain experts. These independent but related properties are therefore crucial ingredients to provide explanations that are accurate and robust, as well as domain-relevant and convincing (Rizzo et al., 2022).

Attribution methods such as Saliency Maps (Simonyan et al., 2013), DeepLIFT (Shrikumar et al., 2017) or Integrated Gradients (Sundararajan et al., 2017) highlight the input features that are deemed important in the decision process as heat maps. These are especially useful, as such maps are easy to interpret and no specific domain knowledge is needed to provide them. Moreover, methods like Integrated Gradients fulfill several desiderata of faithful explanations (Jacovi and Goldberg, 2020), which makes them an even more attractive option to explain DNNs.

Code for DARE can be found at: https://github.com/ibm/domain-adaptive-attribution-robustness

However, recent work has shown that attribution maps do not necessarily fulfill the *robustness* aspect of faithful explanations. In particular, it has been shown in the vision domain that small input alterations can be crafted so as to change the attribution maps drastically, while leaving the prediction unchanged (Ghorbani et al., 2019). Very recently, the same phenomenon has been confirmed in the textual domain as well (Ivankay et al., 2022a).

But what about the plausibility of the explanations? This paper starts by first pointing out that the importance of plausibility has been overlooked in favor of exclusively focusing on faithfulness, in particular in the textual domain. This is important because, when trying to protect a system from adversarial attacks against explanations, it is not only crucial to quantify their faithfulness, but also the plausibility of the possible adversarial samples. In fact, unfaithful but plausible explanation attacks — *convincing lies* — have been pointed out to be particularly pernicious, since they are more difficult for domain experts to spot than equally unfaithful and implausible perturbations and explanations — *unconvincing lies* — (Rizzo et al., 2022). These observations are crucial for the use of AI explainability in high stakes scenarios, as in automated medical diagnosis, EHR classification or triage (Girardi et al., 2018), a medical professional might overlook some critical areas in a cancer cell image or disregard certain important words because they do not appear relevant according to an otherwise *plausible explanation*.

In this work, we focus on the robustness aspect of faithful and plausible attributions in biomedical text classification problems. Specifically, we investigate how to extract in-context adversarial perturbations which are plausible in each specific task domain under consideration. Then, we propose our attribution robustness (AR) estimator that quantifies AR in a domain-specific way. Finally, we explore methods to mitigate the domain-specific fragility of explanation methods in order to train text classifiers that can safely be deployed in safety-critical use cases like healthcare. We summarize our contributions as follows:

- We conceptually relate faithfulness and plausibility to domain-specific attribution robustness estimation on textual data.
- To this end, we extend previous work on AR estimation and introduce our AR estimator, DOMAINADAPTIVEARESTIMATOR (DARE), based on domain-plausible attacks that can be used to estimate AR in a domain-specific way.
- We then empirically show that attribution maps are susceptible to adversarial perturbations that are plausible in the biomedical domain on three (multilabel) medical datasets.
- We are the first to develop and empirically validate two methods to mitigate adversarial perturbations and train text classifiers with robust attribution methods.

2 Related work

Recently, attribution methods like Saliency Maps (Simonyan et al., 2013), Integrated Gradients (Sundararajan et al., 2017), DeepLIFT (Shrikumar et al., 2017) or Shapley Values (Lundberg and Lee, 2017) have been widely deployed in the medical domain for deep learning (Tjoa and Guan, 2020). These methods aim to provide insight into the inference process in DNNs. They highlight features in the input that are deemed relevant in the decision process, without requiring any domain-specific knowledge or heavy computation resources. Thus, they have been widely adopted in areas where predictions need to be accompanied by explanations, such as analysis of medical images (Zhao et al., 2018; Arras et al., 2017) or important symptoms that contribute to or against a given diagnosis (Ribeiro et al., 2016; Tjoa and Guan, 2020). The authors Tjoa and Guan (2020) provide an extensive survey on such methods applied in several medical problem setups.

The work of Ghorbani et al. (2019) explore the robustness of such attribution methods and find that they are susceptible to adversarial perturbations, both in the image domain (Dombrowski et al., 2019; Ivankay et al., 2021), and the text domain (Ivankay et al., 2022a; Sinha et al., 2021; Atmakuri et al., 2022; Ivankay et al., 2022b). However, these works operate on general, non-domain-specific text. None investigate domain-specific text, such as healthcare, where most datasets possess unique vocabularies and semantics. We aim to provide insight into how current methods can be adapted to such specific technical domains.

In order to mitigate the highlighted fragility of attributions in DNNs, several methods have been developed. The authors Chen et al. (2019); Dombrowski et al. (2019); Singh et al. (2019) propose

	VANILLA	ADVERSARIAL	FAR-IG
Original	'took zoloft for 5 months. no side effects except sexual dysfunction. i didn't feel much better or happier and it made me feel really drowsy.' $F(\boldsymbol{s}, l =$"4.0") = 1.0	'took zoloft for 5 months. no side effects except sexual dysfunction. i didn't feel much better or happier and it made me feel really drowsy.' $F(\boldsymbol{s}, l =$"4.0") = 1.0	'took zoloft for 5 months. no side effects except sexual dysfunction. i didn't feel much better or happier and it made me feel really drowsy.' $F(\boldsymbol{s}, l =$"4.0") = 1.0
Adversarial	'took zoloft for 5 months. no side effects except sexual dysfunction. i didn't feel much better or anything and it made me feel really drowsy.' $F(\boldsymbol{s}_{\text{adv}}, l =$"4.0") = 1.0 **Cos.** = -0.32 $MedSTS$ = 0.99	'took zoloft for 5 months. no side effects except sexual dysfunction. i didn't feel much better or stronger and it made me feel really drowsy.' $F(\boldsymbol{s}_{\text{adv}}, l =$"4.0") = 1.0 **Cos.** = -0.05 $MedSTS$ = 0.96	'took zoloft for 5 months. no side effects except nerve dysfunction. i didn't feel much better or happier and it made me feel really drowsy.' $F(\boldsymbol{s}_{\text{adv}}, l =$"4.0") = 1.0 **Cos.** = 0.79 $MedSTS$ = 0.93

Table 1: Attribution methods in medical text classifiers on the Drug Reviews dataset (Gräßer et al., 2018), trained without any robust objectives (VANILLA) are susceptible to imperceptible word substitutions. By changing one words in the original sample (underlined), the words with originally positive attributions (red) are assigned negative values (blue), and vice versa, while keeping the prediction confidence F in the correct class unchanged. This is indicated by the *Cosine Distance* (Cos.) between the explanations of original and adversarial samples. Attacks on attributions in networks trained with robust training objectives (ADVERSARIAL and our novel FAR-IG) are less successful (higher Cos. values) while also being more perceptible (lower medical semantic similarity - MedSTS - values between original and adversarial samples).

methods that smoothen the decision boundary of the classifiers, making gradients smoother as well. The work of Ivankay et al. (2021) provides a general framework to perform adversarial training of attributions, successfully making attributions more robust to input perturbations. However, all of these methods have been developed for the continuous image domain. The transition of such methods to the discrete input space like text has not been investigated, nor has any novel method for text been introduced. In this work, we demonstrate how these shortcomings can be mitigated.

3 Background and Motivation

In this chapter, we introduce the background and motivation of our AR estimation. We define a text classifier F as a function that maps a text sample $\boldsymbol{s}$ to a label l from a given set of labels $\mathbb{L}$. In discrete input spaces like text, F is a function composition of a non-differentiable embedding E that maps the discrete inputs into a continuous domain $\mathbb{R}^{h \times p}$, and a differentiable classifier f mapping the embeddings to the output logits $\mathbb{R}^{|\mathbb{L}|}$. We denote $\mathbb{S} = \{(\boldsymbol{s}, l) \,|\, \boldsymbol{s} = (w_i)_{i \in \{0 \ldots |\boldsymbol{s}|-1\}},\ w_i \in \mathbb{W},\ l \in \mathbb{L},\ |\mathbb{S}| = N\}$ as the set of N text samples $\boldsymbol{s}$ with a label $l \in \mathbb{L}$, each containing a sequence of words w_i drawn from the vocabulary $\mathbb{W}$, h the embedding dimension and p the maximum sequence length. *Attributions* are functions $\boldsymbol{a} = A(\boldsymbol{s}, F, l)$ that assign a value to each word w_i in a text sample $\boldsymbol{s}$, indicating its importance in the DNN inference process. We sum up the attribution values of each w_i-s embedding, resulting in a single value for each word.

Attribution robustness is defined as the Lipschitz attribution robustness constant (Ivankay et al., 2022b), given in the following equation:

$$r(\boldsymbol{s}) = \max_{\tilde{\boldsymbol{s}} \in \mathcal{N}(\boldsymbol{s})} \frac{d\big[A(\tilde{\boldsymbol{s}}, F, l),\ A(\boldsymbol{s}, F, l)\big]}{d_s(\tilde{\boldsymbol{s}}, \boldsymbol{s})} \tag{1}$$

with the prediction constraint

$$\arg\max_{i \in \{1 \ldots |\mathbb{L}|\}} F_i(\tilde{\boldsymbol{s}}) = \arg\max_{i \in \{1 \ldots |\mathbb{L}|\}} F_i(\boldsymbol{s}) \tag{2}$$

Here, $r(\boldsymbol{s})$ denotes the robustness of attribution method A computed for text sample $\boldsymbol{s}$ with label l, drawn from $\mathbb{L}$, and classifier F. The function $d[A(\tilde{\boldsymbol{s}}, F, l), A(\boldsymbol{s}, F, l)]$ denotes the distance between original and adversarial attribution maps $A(\tilde{\boldsymbol{s}}, F, l)$ and $A(\boldsymbol{s}, F, l)$, $\mathcal{N}(\boldsymbol{s})$ is a predefined neighborhood of text sample $\boldsymbol{s}$. The term $d_s(\tilde{\boldsymbol{s}}, \boldsymbol{s})$ indicates the distance of adversarial and original input texts. The robustness of an attribution method on a test dataset $\mathbb{S}$ then becomes the average of $r(\boldsymbol{s})$ over the dataset. Note that the robustness of an attribution method on a classifier is inversely proportional to the constant computed in Equation (1), as large attribution distances and small input distances result in large constants, indicating low robustness. This reflects the definition of the robustness property of *faithful* explanations (Jacovi and Goldberg, 2020).

Our first contribution is conceptual and is motivated by the observation that *plausibility* is a criterion that is rooted in the specific domain and

the semantic conventions within it. Thus, methods to guard against adversarial attacks on explanations need to be *domain adaptive* to conform to the threat model that prioritizes unfaithful explanations which are *semantically plausible* in the domain under consideration, thus are particularly misleading and potentially dangerous.

The strategy we propose to control *domain-adaptive plausibility* is based on the observation that while the numerator in Equation (1) characterizes *faithfulness* by quantifying the effect of adversarial attacks on attributions, the denominator can be adapted to capture *plausibility* by promoting adversarial attacks that remain close to the original input in a semantically meaningful way, in the domain under consideration. In particular, while Ivankay et al. (2022a) utilize the cosine distance of sentence embeddings obtained from domain-agnostic encoders like Universal Sentence Encoder (Cer et al., 2018) and MiniLM (Wang et al., 2020a), we can obtain a domain-specific measure of distance by using embeddings trained on the domain of interest. This will control the plausibility of the adversarial samples by making sure that their domain-dependent semantic distance remains close to the original inputs.

Table 1 exemplifies this approach of quantifying the fragility of attributions in medical text by simultaneously keeping track of faithfulness and plausibility through domain-adapted semantic similarity.

4 Medical Attribution Robustness

Current AR estimation algorithms (Ivankay et al., 2022a; Sinha et al., 2021) were designed to operate in the general text domain, such as news articles (Zhang et al., 2015; Lifferth, 2018), movie reviews (Maas et al., 2011) or product reviews (Asghar, 2016) and make use of the generously available labeled data in these domains. This section describes our proposed methods to adapt these algorithms to the biomedical and healthcare domains where data is sparse and the vocabularies are domain-specific. We describe our datasets and models, we observe that current estimators can be made domain-adaptive by abstracting the candidate extractor and finally, inspired by the works of Ivankay et al. (2022b), we introduce our estimation algorithm DOMAINADAPTIVEARESTIMATOR (DARE), which can effectively be used to estimate AR in the domain of biomedical text.

4.1 Medical datasets

In healthcare, text can appear in many different forms with diverse vocabularies. Thus, we choose three text datasets that cover different aspects of relevant use cases in the medical domain. Often, the datasets are not large enough to train models with state-of-the-art numbers of parameters, such as transformers. Therefore, we make heavy use of transfer learning by utilizing pretrained transformer-based language models and finetune them on our datasets.

Our first dataset, Drug Reviews (DR) (Gräßer et al., 2018), consists of patient reviews of different medical drugs, classified into a rating of 1 to 10 for patient satisfaction. The dataset contains 215063 samples, written in mostly layman's terms along with the names of the drugs and symptom descriptions. Given the dataset's nature, the classification model we choose is a finetuned RoBERTa model, with pretrained weights from Hugging Face (Wolf et al., 2020).

The Hallmarks of Cancer (Baker et al., 2016) dataset (HoC) consists of 1852 biomedical publication abstract associated with 0 or more hallmarks of cancer (Hanahan, 2022). The samples are peer-reviewed publication texts, containing few to no misspellings with scientific biomedical vocabulary. As the dataset contains only a small amount of samples, we finetune a pretrained BioLinkBERT (Yasunaga et al., 2022) model from Hugging Face to achieve state-of-the-art classification accuracy on this dataset.

Lastly, we evaluate the MIMIC-III (Johnson et al., 2016) Discharge Summary dataset (MIMIC). This is a set of extremely long, de-identified, free text ICU discharge summaries from patients admitted to critical care, written by medical professionals. The corresponding ICD-9 codes (World Health Organization, 1988) are associated with each sample in a multilabel fashion. This dataset contains in average 2500 words per sample (Johnson et al., 2016), thus traditional BERT-based models are not feasible as their runtime scales quadratically with the sequence length. Therefore, we finetune a pretrained Clinical-Longformer model (Li et al., 2022), a Longformer MLM (Beltagy et al., 2020) trained on the MIMIC-III discharge summaries. For an in-depth, more detailed description of our datasets and models, we refer to Appendix A.1.

Algorithm 1 `DomainAdaptiveAREstimator`

Input: Input $\boldsymbol{s}$ with label set l, classifier F, attribution A, distance metric d, prediction constraint P, language model MLM, number of candidates $|\mathbb{C}|$, maximum perturbation word ratio ρ_{max}

Output: Adversarial sentence $\boldsymbol{s}_{\text{adv}}$

$\boldsymbol{s}_{\text{adv}} \leftarrow \boldsymbol{s}$, $d_{max} \leftarrow 0$, $n \leftarrow 0$
$I_{\boldsymbol{s}} = \nabla_{\boldsymbol{s}} d\big[A(\boldsymbol{s}+\epsilon, F, l),\ A(\boldsymbol{s}, F, l)\big]$
for $w_i \in \langle w_1, ..., w_{|\boldsymbol{s}|}\rangle | I_{m-1} \geq I_m \forall m \in \{2, ..., |\boldsymbol{s}|\}$ **do**
  if $w_i \in \mathbb{S}_{\texttt{Stop words}}$ **then**
    continue
  $\mathbb{C}_i \leftarrow \text{MLM}(w_i, \boldsymbol{s}, |\mathbb{C}|)$
  for $c_k \in \mathbb{C}_i$ **do**
    $\tilde{\boldsymbol{s}}_{w_{ik}} \leftarrow$ Replace w_i in $\boldsymbol{s}_{\text{adv}}$ with c_k
    if $\text{P}(F_j(\tilde{\boldsymbol{s}}_{w_{ik}}), l)$ not satisfied **then**
      continue
    $\tilde{d} = d\big[A(\tilde{\boldsymbol{s}}_{w_{ik}}, F, l), A(\boldsymbol{s}, F, l)\big]$
    if $\tilde{d} > d_{max}$ **then**
      $\boldsymbol{s}_{\text{adv}} \leftarrow \tilde{\boldsymbol{s}}_{w_{ik}}$
      $d_{max} \leftarrow \tilde{d}$
      $n \leftarrow n + 1$
    if $\rho = \frac{n+1}{|\boldsymbol{s}|} > \rho_{max}$ **then**
      break

4.2 AR in multilabel datasets

Many text classification datasets in healthcare do not only have one label per sample. In HoC, multiple hallmarks can be associated with an abstract, and MIMIC contains hardly any discharge summary with only one associated ICD-9 code. However, current AR estimation definitions only focus on the single label case. Therefore, we make the following modifications to make AR work in the multilabel case. First, we modify the prediction constraint from Equation (1) to reflect multilabel predictions. The label l becomes a set of predicted labels, and the prediction constraint in Equation (2) holds as long as the predicted set of labels from the original sample is equal to the one from the adversarial sample. We denote this constraint as P in our estimation algorithm. Second, attribution methods compute maps on a per-class basis, where the overall attribution $\mathbf{A} = A(\boldsymbol{s}, F, l)$ equals the attribution of the single predicted class l. In a multilabel case, we extend this notion to the sum of attributions for each predicted class, thus the overall attribution map becomes $\mathbf{A} = \sum_{l_i \in l} A(\boldsymbol{s}, F, l_i)$.

4.3 DomainAdaptiveAREstimator (DARE)

Candidate extractors are essential parts of AR estimators, as they provide substitution candidates for the input words, largely contributing to the plausibility and perceptibility of the adversarial alterations. We find that candidate extractors in current work (Ivankay et al., 2022a,b), the counter-fitted synonym embeddings (Mrkšic et al., 2016) and the masked language model (MLM) DistilBERT (Sanh et al., 2019a), are suboptimal in our case, due to their vocabulary only minimally overlapping with the ones from our datasets. However, following the idea of Ivankay et al. (2022b), we argue that, when using the right model, MLMs are in fact effective candidate extractors for word substitutions. Not only do they take context of the words into account, but can be trained on unlabeled data in an unsupervised fashion, thus pretrained models are available for many domains and use cases. Therefore, they can easily be adapted to any domain, without the need for labeled synonym data. For this reason, as our substitution candidate extractors, we choose a pretrained MLM that maximizes the top-5 accuracy of predicting the words in dataset, when each is masked separately, averaged over the dataset. This metric is used as it represents how well the MLMs capture the context of the words, providing meaningful and in-context substitution candidates that will likely result in fluent adversarial samples. Consequently, we use the MLMs DistilRoBERTa (Sanh et al., 2019b) for Drug Reviews, PubMedBERT (Gu et al., 2021) for HoC and Clinical-Longformer (Li et al., 2022) for MIMIC-III. Table 2 summarizes the accuracies of the MLMs that we have tested.

In order to estimate the AR of our classifiers, we propose our two-step, domain-adaptive AR estimator, DARE, written in Algorithm 1. In the first step, an importance ranking of the words in the text samples is extracted in order to prioritize words that are *likely* to impact attributions when substituted. In contrast to current work, we use the gradient of attribution distance as ranking, as this is computationally less heavy than substituting each word with the mask token and performing a single forward pass for each. The second step of DARE is then the extraction of in-context candidates for the highest ranked words, with the pretrained MLMs discussed above and substituting the words greedily with the candidate that maximizes $r(\boldsymbol{s})$ in Equation (1). This allows for efficiently characterizing the robustness aspect of faithfulness while making

MLM	HoC	Drug Reviews	MIMIC-III
BERT (Kenton and Toutanova, 2019)	0.786	0.702	0.677
DistilBERT (Sanh et al., 2019a)	0.733	0.599	0.580
DistilRoBERTa (Sanh et al., 2019b)	0.768	**0.745**	0.604
PubMedBERT (Gu et al., 2021)	**0.908**	0.704	0.781
BioClinicalBERT (Alsentzer et al., 2019)	0.775	0.629	0.847
ClinicalBigBird (Li et al., 2022)	-	-	0.372
Clinical-Longformer (Li et al., 2022)	-	-	**0.867**

Table 2: Top-5 accuracies of the masked language models (MLMs) on our datasets Hallmarks of Cancer (HoC), Drug Reviews and MIMIC-III. Each word in each sample of the dataset is masked and the sample is then propagated through the MLM. If the original masked word is in the top-5 predictions of the MLM, the sample counts as positive.

sure the substitutions are in-context, relevant and maintain the plausibility of attributions.

5 Robust Attributions

In this section, we describe our methods to mitigate fragility of attribution maps in text. Specifically, we are the first to introduce adversarial training (Madry et al., 2018) as a baseline (Sinha et al., 2021) and our adapted FAR (Ivankay et al., 2021) training as a novel method to achieve state-of-the-art attribution robustness in deep neural networks for text classification. Even though we describe and later evaluate the methods on biomedical datasets, these are general training methods that are applicable to any text classification problem.

5.1 Adversarial Training

In an untargeted setup, adversarial training (Moosavi-Dezfooli et al., 2016; Madry et al., 2018) augments the training data with samples $\boldsymbol{s}_{\text{adv}}$ specifically computed as a function of $\boldsymbol{s}$ to maximize the classification loss l_c, written in Equation (3).

$$\boldsymbol{s}_{\text{adv}} = \underset{\tilde{\boldsymbol{s}} \in \mathcal{N}(\boldsymbol{s})}{\arg\max}\, l_c(\tilde{\boldsymbol{s}}, F, l) \tag{3}$$

where $\mathcal{N}$ denotes the search neighborhood of original sample $\boldsymbol{s}$, F the classifier and l the true label of sample $\boldsymbol{s}$. The classifiers then are trained following the optimization objective in Equation (4).

$$\theta^* = \underset{\theta}{\arg\min} \sum_{\boldsymbol{s} \in \mathbb{S}} l_c(\boldsymbol{s}_{\text{adv}}, F, l) \tag{4}$$

where θ^* denotes the optimal model parameters. It has been shown both in the image (Singh et al., 2019; Dombrowski et al., 2019; Chen et al., 2019) and the text domain (Sinha et al., 2021) that adversarial training not only enhances prediction robustness in classifiers, but also improves attribution robustness.

In order to solve the inner optimization problem in Equation (4), we choose the A2T (Yoo and Qi, 2021) attack framework, as it provides flexibility in terms of candidate extraction methods and is optimized for adversarial training runtime. By adapting A2T to use our the MLMs described in Section 4.3, we successfully extract in-context and imperceptible adversarial samples for training.

5.2 FAR for Text

The authors Ivankay et al. (2021) introduced a general framework for training robust attributions (FAR) in deep neural networks in the image domain. They achieve state-of-the-art robustness with few assumptions about the networks or attribution methods. Intuitively, FAR performs adversarial training on attributions and trains networks to minimize the maximal distance between original and adversarial attributions. Equation (5) describes their extraction of adversarial samples for training.

$$\begin{aligned}\boldsymbol{s}_{\text{adv}} = \underset{\tilde{\boldsymbol{s}} \in \mathcal{N}(\boldsymbol{s})}{\arg\max} \Big\{ & (1-\gamma) \cdot l_c(\tilde{\boldsymbol{s}}, F, l) \\ & + \gamma \cdot d\big[A(\tilde{\boldsymbol{s}}, F, l),\ A(\boldsymbol{s}, F, l)\big] \Big\}\end{aligned} \tag{5}$$

with $\boldsymbol{s}_{\text{adv}}$ denoting the adversarial sample, $\mathcal{N}$ the neighborhood space of the original sample $\boldsymbol{s}$, l_c the classification loss of classifier F on $\boldsymbol{s}$ with true label l. d denotes a distance between attribution maps A, γ a constant with $0 \leq \gamma \leq 1$.

Given the above extraction of adversarial samples, the authors train robust networks by solving the following optimization in Equation (6).

$$\begin{aligned}\theta^* = \underset{\theta}{\arg\min} \sum_{\boldsymbol{s} \in \mathbb{S}} \Big\{ & (1-\delta) \cdot l_c(\boldsymbol{s}_{\text{adv}}, F, l) \\ & + \delta \cdot d\big[\boldsymbol{s}_{\text{adv}}, F, l),\ A(\boldsymbol{s}, F, l)\big] \Big\}\end{aligned} \tag{6}$$

with the notation kept from the previous sections and δ denoting a constant with $0 \leq \delta \leq 1$.

The algorithm was designed to work in the image domain. It requires each point in the embedding space (pixel space) to be a valid input. In our case, as text is a discrete input space, this does not hold. Thus, to make the method work for text inputs, we make the following adaptations. Instead of extracting the adversarial samples with the gradient-based IFIA algorithm described in the original paper, we utilize our Algorithm 1 from Section 4.3 to solve the inner maximization in Equation (5). To this end, the prediction constraint in Line 9 of DARE (Algorithm 1) can be omitted to allow for adversarial samples that maximize prediction loss. Moreover, the classification loss can be added as an additive term to the attribution loss in Line 11 to enable joint training of robust predictions and attributions. With our modifications, we successfully overcome the drawbacks of FAR while maintaining the benefits of training robust networks.

6 Experiments

In this section, we report our experiments and setup to estimate attribution robustness in the biomedical domain. We compare the robustness of four attribution methods on three text classifiers trained naively and with robust optimization objectives (adversarial training and FAR). Our results show that the naively trained models are heavily sensitive to imperceptible word substitution attacks, while the two robust training methods significantly increase attribution robustness, with FAR outperforming adversarial training.

6.1 Experimental setup

For each dataset described in Section 4.1, we compare the attribution robustness of a classification model trained with three different training objectives: i) a vanilla natural model trained with the cross-entropy loss; ii) a model trained with adversarial training as described in Section 5.1 and iii) a model trained with robust FAR objectives from Section 5.2. The attribution methods evaluated are Saliency (S) (Simonyan et al., 2013), DeepLIFT (DL) (Shrikumar et al., 2017), Integrated Gradients (IG) (Sundararajan et al., 2017) and the models' self-attention weights (A) (Bahdanau et al., 2015). We choose these as they are popular methods to provide explanations for DNNs in healthcare (Tjoa and Guan, 2020). We use DARE from Section 4.3, with the corresponding MLMs from Table 2 to extract adversarial samples and analyze the cosine distance of original and adversarial attributions, the semantic similarity between original and adversarial input text samples (using the MedSTS semantic embeddings) and combining these two metrics, the resulting attribution robustness constants $r(\boldsymbol{s})$, described in Section 3. A complete set of estimation parameters is given in Table 7 of the appendix.

To evaluate the semantic similarity between original and perturbed inputs, current methods utilize state-of-the-art sentence embeddings on the STS-Benchmark dataset (Cer et al., 2017). We argue that this is suboptimal, as it is not clear whether it captures perturbation perceptibility in the biomedical domain as well. Therefore, we utilize the model made public by Yang et al. (2020) to evaluate semantic distance between texts. This model is the top performing RoBERTa model on the MedSTS dataset (Wang et al., 2020b), a state-of-the-art dataset for semantic similarity in the biomedical domain.

Our vanilla (Van.) models are trained with the standard cross-entropy classification loss, the adversarially trained models (Adv.) with the A2T adversarial training framework (Yoo and Qi, 2021), utilizing the MLMs from Table 2 as candidate extractors. To train our FAR robust models (FAR-IG), we use the FAR training framework described in Section 5.2, using DARE to solve the inner maximization of Equation (6), the cosine distance as attribution distance and Integrated Gradients (IG) as attribution method. For reproducibility, we report the full set of training parameters in Table 5, 8 and 9. The estimation is reported with a threefold cross validation, averaging the results. The models and datasets are implemented in PyTorch (Paszke et al., 2019) and PyTorch Lightning (Falcon and The PyTorch Lightning team, 2019), the pretrained weights are taken from the Hugging Face library (Wolf et al., 2020), with the attributions implemented with Captum (Kokhlikyan et al., 2020). The models are finetuned on the datasets using 4 Nvidia A100 GPUs.

6.2 Results

Table 3 summarizes the results of our experiments. We observe that the non-robust vanilla models (Van.) perform poorly in terms of cosine distance between original and adversarial attribution maps compared to their robust counterparts (Adv. and FAR-IG). Especially the attributions DeepLIFT (DL) and Integrated Gradients (IG) are significantly

		$cos(A_{\text{adv}}, A)$				*MedSTS*				$r(\boldsymbol{s})$			
	MODEL	*S*	*DL*	*IG*	*A*	*S*	*DL*	*IG*	*A*	*S*	*DL*	*IG*	*A*
HoC	VAN.	0.67 ±0.22	-0.09 ±0.22	0.06 ±0.27	0.66 ±0.14	0.79 ±0.12	0.79 ±0.13	0.79 ±0.09	0.78 ±0.1	0.76 ±0.11	2.6 ±0.11	2.2 ±0.22	0.77 ±0.11
	ADV.	0.81 ±0.09	0.09 ±0.22	0.46 ±0.23	0.74 ±0.14	0.79 ±0.1	0.79 ±0.13	0.79 ±0.09	0.78 ±0.1	0.45 ±0.11	2.2 ±0.25	1.3 ±0.16	0.59 ±0.09
	FAR-IG	**0.84** ±0.08	**0.24** ±0.2	**0.65** ±0.26	**0.86** ±0.08	0.77 ±0.14	0.77 ±0.14	0.78 ±0.11	0.77 ±0.14	**0.35** ±0.12	**1.6** ±0.31	**0.8** ±0.31	**0.3** ±0.05
Drug Rev.	VAN.	0.89 ±0.12	0.25 ±0.32	0.48 ±0.35	0.72 ±0.18	0.92 ±0.08	0.92 ±0.09	0.92 ±0.09	0.91 ±0.09	0.69 ±0.07	4.1 ±0.19	3.3 ±0.22	2.1 ±0.1
	ADV.	0.91 ±0.12	0.36 ±0.3	0.49 ±0.34	0.78 ±0.17	0.91 ±0.09	0.9 ±0.1	0.91 ±0.09	0.9 ±0.09	0.45 ±0.06	3.7 ±0.17	2.8 ±0.14	1.1 ±0.09
	FAR-IG	**0.93** ±0.11	**0.77** ±0.28	**0.86** ±0.21	**0.86** ±0.12	0.9 ±0.09	0.9 ±0.09	0.9 ±0.09	0.89 ±0.1	**0.35** ±0.05	**1.2** ±0.14	**0.8** ±0.14	**0.73** ±0.07
MIMIC-III	VAN.	0.35 ±0.27	0.08 ±0.33	0.0 ±0.37	0.7 ±0.26	0.88 ±0.07	0.84 ±0.07	0.82 ±0.11	0.84 ±0.07	3.1 ±	2.9 ±0.18	2.8 ±0.15	0.94 ±0.2
	ADV.	**0.44** ±0.32	**0.12** ±0.26	0.0 ±0.45	**0.76** ±0.21	0.85 ±0.07	0.77 ±0.19	0.8 ±0.03	0.81 ±0.13	**1.9** ±0.21	**1.9** ±0.47	**2.5** ±0.27	**0.63** ±0.12
	FAR-IG	–	–	–	–	–	–	–	–	–	–	–	–

Table 3: Attribution robustness metrics (mean and stddev.) of the vanilla (Van.), adversarially trained (Adv.) and FAR-trained (FAR-IG) models, trained on our three datasets. We perform AR estimation for the attributions S, DL, IG and A. The reported metrics are the cosine similarity between attributions of original and adversarial samples - $cos(A_{\text{adv}}, A)$ -, the semantic similarity of the two input text samples - *MedSTS* - as well as the estimated attribution robustness constant - $r(\boldsymbol{s})$ -. We conclude that the vanilla models perform poorly in terms of attribution robustness, while both adversarially and FAR-IG trained models are significantly more robust, yielding higher attribution similarities and lower $r(\boldsymbol{s})$ values. FAR-IG models outperform adversarially trained models, giving the most promising method to train attributionally robust networks.

altered by the attacks. This is reflected in the higher estimated robustness constants $r(\boldsymbol{s})$ for the vanilla models. Thus, we conclude that training networks with no robustness objective is largely suboptimal if faithful and robust explanations are needed.

However, both the baseline adversarial training and our adapted FAR objectives are able to train networks with significantly more robust attributions than vanilla training. For the HoC dataset and IG attributions, adversarial training increases the cosine similarity up to 0.46, while FAR-IG training increases it by 0.65 over. A similar trend is observable for the other models, datasets and attribution methods. FAR-IG training reduces the estimated robustness constants consistently by 40-60%, which is a significant increase in robustness. This convinces us that FAR is a feasible method to achieve robust attributions in DNNs.

We further observe that even if our FAR-IG model is not evaluated on IG, but on S, DL or A, it still outperforms vanilla and adversarially trained models both in terms of $cos(A_{\text{adv}}, A)$ and $r(\boldsymbol{s})$. Therefore, we conclude that the robustness attained by FAR training with IG transfers to other attributions, further strengthening our confidence in FAR being an attractive option to train robust networks.

7 Conclusion

In this work, we explored the attribution robustness of biomedical text classification. We extended current robustness estimators to introduce DARE, a domain-adaptive AR estimator. Then, we showed on three different biomedical datasets that classifiers trained without robust objectives lack robustness to small input perturbations in this domain as well. In order to mitigate this, we proposed two training methods, adversarial training and FAR to train neural networks that yield robust attributions even in the presence of carefully crafted input perturbations. With our experiments, we show that adversarial training and FAR are able to increase the attribution robustness significantly, with FAR giving the best results.

Our work is a key milestone for the deployment of DNNs in the biomedical domain, as such a safety-critical application area requires sound and faithful explanations. In the future, we plan to extend our investigation from text classification to other NLP problems in the biomedical domain. Moreover, investigating the robustness of other

types explanation methods is an important future research direction.

8 Limitations and Risks

DARE only works for text. In its introduced form, it requires the prediction gradients for importance ranking, thus can only be used to attack differentiable architectures (up to the embedding layer). Most state-of-the-art classifiers (DNNs, transformers) fulfill this criteria though. Moreover, DARE requires MLMs trained in a specific domain to work – which might not always be readily available. However, as MLMs can be trained in an unsupervised fashion, pretrained MLMs can be finetuned to that domain with rather low effort.

The main risk of DARE is that it does not give a guaranteed lower bound of robustness. If an attacker develops a stronger attack that is able to compute better perturbations that alter attributions to a greater extent, having a model that is robust to DARE perturbations might not be sufficient to withstand those stronger attacks. Taking the robustness estimation for granted is a risk, as it is true for most other attacks in traditional adversarial setups. This directly indicates another risk, namely that DARE could be used to attack explanations in deployed systems that are not trained robustly.

We train our methods on state-of-the-art Nvidia A100 GPUs. Without having such GPUs available, FAR training in particular becomes a bottleneck, as the computation graph needs to be stored for several forward and backward passes, depending on the attribution method used. On this end, we also require the attributions to be differentiable with respect to the input embeddings, which is an implicit requirement of the FAR training method. We do not see any risks in using FAR to train robust networks.

Finally, we do not examine any other aspects of faithful interpretations, only the robustness. We assume that these methods reflect the model behavior to some extent, but do not conclude any experiments to verify this assumption. Further investigation into whether more robust attributions yield better faithfulness in other aspects could be an interesting future research topic.

References

Emily Alsentzer, John R Murphy, Willie Boag, Wei-Hung Weng, Di Jin, Tristan Naumann, WA Redmond, and Matthew BA McDermott. 2019. Publicly available clinical bert embeddings. *NAACL HLT 2019*, page 72.

Leila Arras, Grégoire Montavon, Klaus-Robert Müller, and Wojciech Samek. 2017. Explaining Recurrent Neural Network Predictions in Sentiment Analysis. In *8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis WASSA 2017: Proceedings of the Workshop*, pages 159–168. The Association for Computational Linguistics.

Nabiha Asghar. 2016. Yelp dataset challenge: Review rating prediction. *arXiv preprint arXiv:1605.05362*.

Shriya Atmakuri, Tejas Chheda, Dinesh Kandula, Nishant Yadav, Taesung Lee, and Hessel Tuinhof. 2022. Robustness of Explanation Methods for NLP Models. *arXiv preprint arXiv:2206.12284*.

Dzmitry Bahdanau, Kyung Hyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. In *3rd International Conference on Learning Representations, ICLR 2015*.

Simon Baker, Anna-Leena Korhonen, and Sampo Pyysalo. 2016. Cancer hallmark text classification using convolutional neural networks.

Iz Beltagy, Matthew E Peters, and Arman Cohan. 2020. Longformer: The Long-Document Transformer. *arXiv preprint arXiv:2004.05150*.

Daniel Cer, Mona Diab, Eneko Agirre, Iñigo Lopez-Gazpio, and Lucia Specia. 2017. SemEval-2017 Task 1: Semantic Textual Similarity Multilingual and Crosslingual Focused Evaluation. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, pages 1–14.

Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Céspedes, Steve Yuan, Chris Tar, et al. 2018. Universal Sentence Encoder. *arXiv preprint arXiv:1803.11175*.

Jiefeng Chen, Xi Wu, Vaibhav Rastogi, Yingyu Liang, and Somesh Jha. 2019. Robust Attribution Regularization. In *Advances in Neural Information Processing Systems*, pages 14300–14310.

Ann-Kathrin Dombrowski, Maximillian Alber, Christopher Anders, Marcel Ackermann, Klaus-Robert Müller, and Pan Kessel. 2019. Explanations can be manipulated and geometry is to blame. In *Advances in Neural Information Processing Systems*, pages 13589–13600.

William Falcon and The PyTorch Lightning team. 2019. PyTorch Lightning.

Amirata Ghorbani, Abubakar Abid, and James Zou. 2019. Interpretation of Neural Networks is Fragile. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 3681–3688.

Ivan Girardi, Pengfei Ji, An-phi Nguyen, Nora Hollenstein, Adam Ivankay, Lorenz Kuhn, Chiara Marchiori, and Ce Zhang. 2018. Patient Risk Assessment and Warning Symptom Detection Using Deep Attention-Based Neural Networks. In *Proceedings of the Ninth International Workshop on Health Text Mining and Information Analysis*, pages 139–148.

Felix Gräßer, Surya Kallumadi, Hagen Malberg, and Sebastian Zaunseder. 2018. Aspect-Based Sentiment Analysis of Drug Reviews Applying Cross-Domain and Cross-Data Learning. DH '18, page 121–125, New York, NY, USA. Association for Computing Machinery.

Yu Gu, Robert Tinn, Hao Cheng, Michael Lucas, Naoto Usuyama, Xiaodong Liu, Tristan Naumann, Jianfeng Gao, and Hoifung Poon. 2021. Domain-Specific Language Model Pretraining for Biomedical Natural Language Processing. *ACM Transactions on Computing for Healthcare (HEALTH)*, 3(1):1–23.

Douglas Hanahan. 2022. Hallmarks of Cancer: New Dimensions. *Cancer Discovery*, 12(1):31–46.

Adam Ivankay, Ivan Girardi, Chiara Marchiori, and Pascal Frossard. 2021. FAR: A General Framework for Attributional Robustness. *The 32nd British Machine Vision Conference, BMVC 2021*.

Adam Ivankay, Ivan Girardi, Chiara Marchiori, and Pascal Frossard. 2022a. Fooling Explanations in Text Classifiers. In *The Tenth International Conference on Learning Representations, ICLR 2022*.

Adam Ivankay, Mattia Rigotti, Ivan Girardi, Chiara Marchiori, and Pascal Frossard. 2022b. Estimating the adversarial robustness of attributions in text with transformers. *arXiv preprint arXiv:2212.09155*.

Alon Jacovi and Yoav Goldberg. 2020. Towards Faithfully Interpretable NLP Systems: How Should We Define and Evaluate Faithfulness? In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4198–4205.

Alon Jacovi and Yoav Goldberg. 2021. Aligning faithful interpretations with their social attribution. *Transactions of the Association for Computational Linguistics*, 9:294–310.

Alistair EW Johnson, Tom J Pollard, Lu Shen, Li-wei H Lehman, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G Mark. 2016. MIMIC-III, a freely accessible critical care database. *Scientific Data*, 3(1):1–9.

Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of NAACL-HLT*, pages 4171–4186.

Narine Kokhlikyan, Vivek Miglani, Miguel Martin, Edward Wang, Bilal Alsallakh, Jonathan Reynolds, Alexander Melnikov, Natalia Kliushkina, Carlos Araya, Siqi Yan, et al. 2020. Captum: A unified and generic model interpretability library for pytorch. *arXiv preprint arXiv:2009.07896*.

Yikuan Li, Ramsey M Wehbe, Faraz S Ahmad, Hanyin Wang, and Yuan Luo. 2022. Clinical-Longformer and Clinical-BigBird: Transformers for long clinical sequences. *arXiv preprint arXiv:2201.11838*.

William Lifferth. 2018. Fake News.

Scott M Lundberg and Su-In Lee. 2017. A Unified Approach to Interpreting Model Predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pages 4768–4777.

Andrew Maas, Raymond E Daly, Peter T Pham, Dan Huang, Andrew Y Ng, and Christopher Potts. 2011. Learning Word Vectors for Sentiment Analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 142–150.

Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. 2018. Towards Deep Learning Models Resistant to Adversarial Attacks. In *International Conference on Learning Representations*.

Seyed-Mohsen Moosavi-Dezfooli, Alhussein Fawzi, and Pascal Frossard. 2016. Deepfool: a simple and accurate method to fool deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2574–2582.

Nikola Mrkšic, Diarmuid OSéaghdha, Blaise Thomson, Milica Gašic, Lina Rojas-Barahona, Pei-Hao Su, David Vandyke, Tsung-Hsien Wen, and Steve Young. 2016. Counter-fitting Word Vectors to Linguistic Constraints. In *Proceedings of NAACL-HLT*, pages 142–148.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, pages 8026–8037.

Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. "Why Should I Trust You?" Explaining the Predictions of any Classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1135–1144.

Mattia Rigotti, Christoph Miksovic, Ioana Giurgiu, Thomas Gschwind, and Paolo Scotton. 2022. Attention-based Interpretability with Concept Transformers. In *International Conference on Learning Representations*.

Matteo Rizzo, Alberto Veneri, Andrea Albarelli, Claudio Lucchese, and Cristina Conati. 2022. A theoretical framework for ai models explainability. *arXiv preprint arXiv:2212.14447*.

Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019a. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.

Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019b. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *ArXiv*, abs/1910.01108.

Avanti Shrikumar, Peyton Greenside, and Anshul Kundaje. 2017. Learning important features through propagating activation differences. In *International Conference on Machine Learning*, pages 3145–3153. PMLR.

Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. 2013. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. *arXiv preprint arXiv:1312.6034*.

Mayank Singh, Nupur Kumari, Puneet Mangla, Abhishek Sinha, Vineeth N Balasubramanian, and Balaji Krishnamurthy. 2019. On the Benefits of Attributional Robustness. *arXiv preprint arXiv:1911.13073*.

Sanchit Sinha, Hanjie Chen, Arshdeep Sekhon, Yangfeng Ji, and Yanjun Qi. 2021. Perturbing Inputs for Fragile Interpretations in Deep Natural Language Processing. In *Proceedings of the Fourth BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP*, pages 420–434.

Mukund Sundararajan, Ankur Taly, and Qiqi Yan. 2017. Axiomatic Attribution for Deep Networks. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70, pages 3319–3328.

E Tjoa and C Guan. 2020. A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI. *IEEE Transactions on Neural Networks and Learning Systems*.

Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. 2020a. MiniLM: Deep Self-Attention Distillation for Task-Agnostic Compression of Pre-Trained Transformers. *Advances in Neural Information Processing Systems*, 33:5776–5788.

Yanshan Wang, Naveed Afzal, Sunyang Fu, Liwei Wang, Feichen Shen, Majid Rastegar-Mojarad, and Hongfang Liu. 2020b. MedSTS: a resource for clinical semantic textual similarity. *Language Resources and Evaluation*, 54(1):57–72.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. 2020. Transformers: State-of-the-Art Natural Language Processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, Online. Association for Computational Linguistics.

World Health Organization. 1988. International classification of diseases—Ninth revision (ICD-9). *Weekly Epidemiological Record= Relevé épidémiologique hebdomadaire*, 63(45):343–344.

X Yang, X He, H Zhang, Y Ma, J Bian, and Y Wu. 2020. Measurement of Semantic Textual Similarity in Clinical Texts: Comparison of Transformer-Based Models. *JMIR Medical Informatics*, 8(11):e19735–e19735.

Michihiro Yasunaga, Jure Leskovec, and Percy Liang. 2022. LinkBERT: Pretraining Language Models with Document Links. In *Association for Computational Linguistics (ACL)*.

Jin Yong Yoo and Yanjun Qi. 2021. Towards Improving Adversarial Training of NLP Models. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 945–956.

Xiang Zhang, Junbo Zhao, and Yann Lecun. 2015. Character-level convolutional networks for text classification. *Advances in Neural Information Processing Systems*, 2015:649–657.

Guannan Zhao, Bo Zhou, Kaiwen Wang, Rui Jiang, and Min Xu. 2018. Respond-CAM: Analyzing Deep Models for 3D Imaging Data by Visualizations. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I*, pages 485–492.

A Appendix

A.1 Models and Datasets

We use three public datasets to evaluate the attribution robustness of biomedical text classifiers. Our main goal is to show how robust attribution methods are on these datasets, thus we do not aim to advance the state-of-the-art for classification accuracy, but train models that achieve close to state-of-the-art performance while being relatively easy to train. For each dataset, we use a 60%-20%-20% split for training, test and validation splits, apply basic preprocessing by lower casing the text, removing characters that are not in the Latin alphabet and remove double spaces, new line symbols and double quotes.

The Drug Reviews (DR) dataset consists of patient reviews of different medical drugs, classified into a rating of 1 to 10 for patient satisfaction. In order to increase classification performance, we reduce the number of classes to 5 by merging classes 1 and 2, 3 and 4, 5 and 6 etc. The dataset contains 215063 samples, and we train a RoBERTa model for classification, with the standard cross entropy loss on the first 128 tokens.

The Hallmarks of Cancer (HoC) dataset comprises 1852 biomedical publication abstract associated with 0 or more hallmarks of cancer, thus is a 10-class multilabel classification dataset. We finetune a pretrained BioLinkBERT model for classification, use the first 256 tokens as inputs to the model after tokenization and utilize the binary cross entropy as classification weight.

Our last dataset, the MIMIC-III Discharge Summary dataset consists of patients' ICU discharge summaries, associated with their ICD-9 codes. In order to reduce the overall number of classes from 1800, we only take the 50 most frequent ICD-9 codes. This results in a total of 59647 samples. As the summaries are very long, we finetune a pretrained Clinical-Longformer model for classification, with a maximum sequence length of 4096, default attention window size and global attention on the `[CLS]` token.

Table 4 summarizes our models, Table 5 contains the used hyperparameters for our finetuning process and Table 6 the resulting accuracies of all our trained models. We use the AdamW optimizer throughout all our experiments.

The Hallmarks of Cancer and Drug Reviews dataset are publicly available datasets. The requirements for MIMIC-III [1] were completed and we comply with their DUA.

PARAMETER	HALLMARKS OF CANCER	DRUG REVIEWS	MIMIC-III
INPUT SHAPE	(256,)	(128,)	(4096,)
NUM. CLASSES	10	5	50
HF MODEL ID	michiyasunaga/BioLinkBERT-base	roberta-base	yikuan8/Clinical-Longformer
NUM. PARAMS	108240394	124649477	148697906

Table 4: Parameters of our classification models.

A.2 AR Estimation and Robust Training

In order to achieve robust attributions, in addition to the vanilla models we train models with robust training objectives. During adversarial training, we augment the training batches with adversarial samples that maximize classification loss. We use the A2T training method for extracting adversarial samples, with the parameters summarized in Table 8. Our FAR models are trained with the robust objectives from Section 5.2, and the hyperparameters are written in Table 9.

[1] https://physionet.org/content/mimiciii/1.4/

PARAMETER	HALLMARKS OF CANCER	DRUG REVIEWS	MIMIC-III
CLASSIFICATION LOSS	Multilabel binary cross entropy	Cross entropy	Multilabel binary cross entropy
LR	0.00001	0.000001	0.00004
BATCH SIZE	128	64	4
EPOCHS	50	50	50
PRECISION	32	32	16
ACCUMULATE GRADIENT BATCHES	1	1	4

Table 5: Parameters used to train our non-robust, vanilla models.

		HALLMARKS OF CANCER			DRUG REVIEWS			MIMIC-III		
	MODEL	*Van.*	*Adv.*	*FAR-IG*	*Van.*	*Adv.*	*FAR-IG*	*Van.*	*Adv.*	*FAR-IG*
NAT.	ACCURACY	0.95	0.94	0.92	0.9	0.92	0.92	0.92	0.9	-
	PRECISION	0.78	0.74	0.62	0.89	0.92	0.92	0.59	0.57	-
	RECALL	0.89	0.82	0.90	0.9	0.92	0.92	0.71	0.61	-
	F1-SCORE	0.82	0.78	0.73	0.9	0.92	0.92	0.64	0.6	-
	LOSS	0.24	0.27	0.27	0.68	0.36	0.32	0.3	0.33	-
ADV.	ACCURACY	0.88	0.89	0.87	0.61	0.67	0.65	0.89	0.9	-
	PRECISION	0.55	0.59	0.5	0.61	0.66	0.65	0.54	0.55	-
	RECALL	0.75	0.7	0.8	0.6	0.67	0.65	0.67	0.61	-
	F1-SCORE	0.61	0.63	0.62	0.6	0.67	0.65	0.59	0.62	-
	LOSS	0.64	0.53	0.44	2.5	1.1	1.2	0.41	0.39	-

Table 6: Natural (**NAT.**) and adversarial (**ADV.**) classification metrics of the non-robust (Van.), adversarially robust (Adv.) and FAR-trained (FAR-IG) models. All metrics are macro-averaged over the samples, as our datasets are highly class-imbalanced.

PARAMETER	HALLMARKS OF CANCER	DRUG REVIEWS	MIMIC-III
CANDIDATE EXTRACTOR	PubMedBERT	DistilRoBERTa	Clinical-Longformer
$\rho_{\max}$	0.05	0.05	0.005
$\lvert\mathbb{C}\rvert$	5	5	3
$d(A_{\text{adv}}, A)$	cosine	cosine	cosine
$d_s(\boldsymbol{s}_{\text{adv}}, \boldsymbol{s})$	MedSTS semantic embeddings	MedSTS semantic embeddings	MedSTS semantic embeddings

Table 7: Hyperparameters used for estimating attribution robustness for our three datasets Hallmarks of Cancer, Drug Reviews and MIMIC-III. Candidate extractor denotes the MLM used for extracting the replacement candidates in DARE, $\rho_{\max}$ the maximum ratio of perturbed words in each sample, $|\mathbb{C}|$ the number of replacement candidates extracted for each word, $d(A_{\text{adv}}, A)$ the attribution distance metric and $d_s(\boldsymbol{s}_{\text{adv}}, \boldsymbol{s})$ the text input distance.

PARAMETER	HALLMARKS OF CANCER	DRUG REVIEWS	MIMIC-III
CANDIDATE EXTRACTOR	PubMedBERT	DistilRoBERTa	Clinical-Longformer
$\rho_{\max}$	0.05	0.05	0.005
$\lvert\mathbb{C}\rvert$	5	5	3
CLASSIFICATION LOSS	Multilabel binary cross entropy	Cross entropy	Multilabel binary cross entropy
RATIO OF ATTACKED SAMPLES IN BATCH	0.3	0.3	0.3
LR	0.00001	0.000001	0.000001
BATCH SIZE	32	64	16
EPOCHS	30	20	20

Table 8: Parameters used to train our adversarially robust networks.

PARAMETER	HALLMARKS OF CANCER	DRUG REVIEWS
CANDIDATE EXTRACTOR	PubMedBERT	DistilRoBERTa
$\rho_{\max}$	0.05	0.05
A	IG	IG
$d(A_{\text{adv}}, A)$	cosine	cosine
$\lvert\mathbb{C}\rvert$	5	5
CLASSIFICATION LOSS	Multilabel binary cross entropy	Cross entropy
FAR INSTANTIATION	AdvAAT	AAT
γ	0.85	0.0
δ	0.85	0.7
LR	0.00001	0.000001
BATCH SIZE	4	8
EPOCHS	30	20
RATIO OF ATTACKED SAMPLES IN BATCH	0.6	0.6

Table 9: Parameters used to train our FAR-IG networks.

A.3 More Examples

	VANILLA	ADVERSARIAL	FAR-IG
Original	'i have been on invokana since september 2013, so a little over a year. i have experienced hair loss, tiredness, and yeast infections. i talked to my doctor about the hair loss, which i experienced for over a year. he has upped my metformin to the maximum dosage. my hair has stopped falling out. i am also using rosemary essential oil to help with hair loss, and probiotics for the yeast infection. i have had amazing results with this medication in regards to blood sugar control. my a1c went from 12.3 to 7.1 i have never had $F(\boldsymbol{s}, l =$"8.0") = 1.0	'i have been on invokana since september 2013, so a little over a year. i have experienced hair loss, tiredness, and yeast infections. i talked to my doctor about the hair loss, which i experienced for over a year. he has upped my metformin to the maximum dosage. my hair has stopped falling out. i am also using rosemary essential oil to help with hair loss, and probiotics for the yeast infection. i have had amazing results with this medication in regards to blood sugar control. my a1c went from 12.3 to 7.1 i have never had $F(\boldsymbol{s}, l =$"8.0") = 0.88	'i have been on invokana since september 2013, so a little over a year. i have experienced hair loss, tiredness, and yeast infections. i talked to my doctor about the hair loss, which i experienced for over a year. he has upped my metformin to the maximum dosage. my hair has stopped falling out. i am also using rosemary essential oil to help with hair loss, and probiotics for the yeast infection. i have had amazing results with this medication in regards to blood sugar control. my a1c went from 12.3 to 7.1 i have never had $F(\boldsymbol{s}, l =$"8.0") = 0.97
Adversarial	'i have been on invokana since september 2013, so a little over a year. i have noticed scalp loss, tiredness, and yeast infections. i talked to my doctor about the hair loss, which i experienced for over a year. he has upped my metformin to the maximum dosage. my hair has stopped falling out. i am also using rosemary essential oil to help with hair loss, and probiotics for the yeast infection. i have had numerous results with this medication in regards to blood glucose control. my a1c went from 12.3 to 7.1 i have never had $F(\boldsymbol{s}_{\text{adv}}, l =$"8.0") = 0.77 **Cos.** = -0.07 $MedSTS$ = 1.0	'i have been on invokana since september 2013, taking a tad over a year. i have experienced hair loss, tiredness, and yeast infections. i complained to my doctor about the hair loss, which i experienced for over a year. he has upped my metformin to the recommended dosage. my hair has stopped falling out. i am also using rosemary essential oil to help with hair loss, and probiotics for the yeast infection. i have had amazing results with this medication in regards to blood sugar control. my a1c went from 12.3 to 7.1 i have never had $F(\boldsymbol{s}_{\text{adv}}, l =$"8.0") = 0.73 **Cos.** = 0.35 $MedSTS$ = 1.0	'i have been on invokana since september 2013, so a little over a year. i have experienced hair loss, tiredness, and yeast infections. i talked to my doctor about the hair loss, which i experienced for over a year. he has upped my metformin to the recommended dosage. my hair has stopped falling out. i am also using rosemary olive oil to help with hair loss, and probiotics for the yeast infection. i have had amazing success with this medication in regards to blood pressure control. my a1c went from 12.3 to 7.1 i have never had $F(\boldsymbol{s}_{\text{adv}}, l =$"8.0") = 0.93 **Cos.** = 0.58 $MedSTS$ = 1.0
Original	'i have had intractable migraine for 28 years, and migraines from the age of 10 until 28 years ago when it never quit.. i went through many different trials of treatment & nothing worked, so finally the headache specialist gave me vicodin.it worked and i was able to begin living life again. then a new md took the vicodin away and gave me topamax. my life was hell. i live alone in a 2 story house and i had to scoot up/down on my butt. i am 66 & disabled (from strokes) and i was terrified. $F(\boldsymbol{s}, l =$"2.0") = 1.0	'i have had intractable migraine for 28 years, and migraines from the age of 10 until 28 years ago when it never quit.. i went through many different trials of treatment & nothing worked, so finally the headache specialist gave me vicodin.it worked and i was able to begin living life again. then a new md took the vicodin away and gave me topamax. my life was hell. i live alone in a 2 story house and i had to scoot up/down on my butt. i am 66 & disabled (from strokes) and i was terrified. $F(\boldsymbol{s}, l =$"2.0") = 0.98	'i have had intractable migraine for 28 years, and migraines from the age of 10 until 28 years ago when it never quit.. i went through many different trials of treatment & nothing worked, so finally the headache specialist gave me vicodin.it worked and i was able to begin living life again. then a new md took the vicodin away and gave me topamax. my life was hell. i live alone in a 2 story house and i had to scoot up/down on my butt. i am 66 & disabled (from strokes) and i was terrified. $F(\boldsymbol{s}, l =$"2.0") = 0.99
Adversarial	'i have had intractable migraine for 28 years, and migraines from the age of 10 until 28 years old when it never quit.. i went through many different trials of treatment & nothing worked, so finally the arthritis specialist gave me vicodin.it worked and i was able to begin living life again. then a new md took the vicodin away and gave me topamax. my life was saying i was alone in a 2 story house and i had to scoot up/down on my butt. i am 66 & disabled (from strokes) and i was terrified. $F(\boldsymbol{s}_{\text{adv}}, l =$"2.0") = 0.71 **Cos.** = 0.01 $MedSTS$ = 0.89	'i have had intractable epilepsy for 28 years, and migraines from the age of 10 until 28 years ago when it never quit.. i went through many different trials of treatment & nothing worked, so finally the epilepsy specialist gave me vicodin.it worked and i was able to begin living life again. then a new md took the vicodin away and gave me topamax. my life was ruined i live alone in a 2 nd house and i had to scoot up/down on my butt. i am 66 & disabled (from strokes) and i was terrified. $F(\boldsymbol{s}_{\text{adv}}, l =$"2.0") = 0.98 **Cos.** = 0.26 $MedSTS$ = 0.81	'i have had intractable epilepsy for 28 years, and migraines from the age of 10 until 28 years ago when it never quit.. i went through many different trials of treatment & nothing worked, so finally the epilepsy specialist gave me vicodin.it worked and i was able to begin living life again. then a new md took the vicodin away and gave me topamax. my life was hell. i live alone in a 2 nd flat and i had to scoot up/down on my butt. i am 66 & disabled (from strokes) and i was terrified $F(\boldsymbol{s}_{\text{adv}}, l =$"2.0") = 0.99 **Cos.** = 0.58 $MedSTS$ = 0.81
Original	'i have only been using nuva ring for 5 days... i have not been sick in any way.. or had mood swings.. ive noticed i have alittle nore energy to get things done around the house. my sex drive i believe has increased a tiny bit... already was high but i haven't had sex yet since i have had it in due to my partners work schedule.i do feel blowed everyday n i get pains in my stomach here and their like period cramps but nothing to intence. i so far do really like this birth control.. i hope it makes my period leas painful and... $F(\boldsymbol{s}, l =$"8.0") = 1.0	'i have only been using nuva ring for 5 days... i have not been sick in any way.. or had mood swings.. ive noticed i have alittle nore energy to get things done around the house. my sex drive i believe has increased a tiny bit... already was high but i haven't had sex yet since i have had it in due to my partners work schedule.i do feel blowed everyday n i get pains in my stomach here and their like period cramps but nothing to intence. i so far do really like this birth control.. i hope it makes my period leas painful and... $F(\boldsymbol{s}, l =$"8.0") = 0.93	'i have only been using nuva ring for 5 days... i have not been sick in any way.. or had mood swings.. ive noticed i have alittle nore energy to get things done around the house. my sex drive i believe has increased a tiny bit... already was high but i haven't had sex yet since i have had it in due to my partners work schedule.i do feel blowed everyday n i get pains in my stomach here and their like period cramps but nothing to intence. i so far do really like this birth control.. i hope it makes my period leas painful and... $F(\boldsymbol{s}, l =$"8.0") = 0.91

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	'i have only been using nuva ring for 5 days... i have not been manic in any way.. or had mood swings.. ive noticed i have alittle nore energy to get things done around the house. my sex drive i believe has increased a tiny bit... already was high but i haven't had sex yet since i have had it in due to my partners work schedule.i do feel blowed everyday n i get pains in my stomach here and their like period cramps but nothing to intence. lol so NOT do i like this under control.. i hope it makes my period leas painful and... $F(\boldsymbol{s}_{\text{adv}}, l =$"8.0") = 1.0 **Cos.** = -0.35 $MedSTS$ = 0.91	'i have only been using nuva ring for 5 days... i have not been manic in any way.. or had mood swings.. ive noticed i have alittle nore energy to get things done around the house. my sex drive i believe has increased a tiny bit... already was high but i haven't had sex yet since i have had it in due to my partners work schedule.i do feel blowed everyday n i get pains in my stomach here and their like period cramps but nothing to intence. i so i would not take this birth control.. i hope it makes my period leas painful and... $F(\boldsymbol{s}_{\text{adv}}, l =$"8.0") = 0.96 **Cos.** = -0.12 $MedSTS$ = 0.91	'i have only been using nuva ring for 5 days... i have not been depressed in any way.. or had mood swings.. ive glad i have alittle nore energy to get things done around the house. my gas density i believe has increased a tiny bit... already was high but i haven't had intercourse yet since i have had it in due to my partners work schedule.i do feel blowed everyday n i get pains in my stomach here and their like period cramps but nothing to intence. i so far do really like this birth control.. i hope it makes my period leas painful and... $F(\boldsymbol{s}_{\text{adv}}, l =$"8.0") = 0.48 **Cos.** = 0.59 $MedSTS$ = 0.73
Original	'i started using this product a little more than a week ago.i applied it three nights in a row as instructed, and went to a party the next day to test it out.i still sweated, but not nearly as much, and i had hope that with time i would be totally sweat free.i applied it once again the following night, only to continue to sweat the next day.since then (it's been about four days) i have applied hypercare every night without any improvements in the amount i sweat.today was the first day of school and i was sweat the entire day, unable to lift my $F(\boldsymbol{s}, l =$"2.0") = 1.0	'i started using this product a little more than a week ago.i applied it three nights in a row as instructed, and went to a party the next day to test it out.i still sweated, but not nearly as much, and i had hope that with time i would be totally sweat free.i applied it once again the following night, only to continue to sweat the next day.since then (it's been about four days) i have applied hypercare every night without any improvements in the amount i sweat.today was the first day of school and i was sweat the entire day, unable to lift my $F(\boldsymbol{s}, l =$"2.0") = 0.97	'i started using this product a little more than a week ago.i applied it three nights in a row as instructed, and went to a party the next day to test it out.i still sweated, but not nearly as much, and i had hope that with time i would be totally sweat free.i applied it once again the following night, only to continue to sweat the next day.since then (it's been about four days) i have applied hypercare every night without any improvements in the amount i sweat.today was the first day of school and i was sweat the entire day, unable to lift my $F(\boldsymbol{s}, l =$"2.0") = 1.0
Adversarial	'i started using this product a little more than a week ago.i applied it three nights in a row as instructed, and went to a party the next day to test it out.i still sweated, but not nearly as much, and i was hope that by time i would be totally sweat free.i applied it once again the following night, only to continue to sweat the next day.since then (it's been about four days) i have applied this every night without any breaks in the night i sweat.today was the first day of school and i was sweat the entire day, unable to lift my $F(\boldsymbol{s}_{\text{adv}}, l =$"2.0") = 0.87 **Cos.** = -0.27 $MedSTS$ = 1.0	'i started using this product a little more than a week ago.i applied it three nights in a row as instructed, and went to a party the next day to test it out.i still sweated, but not nearly as much, and i had thought that with time i would be totally much less applied it once again the following night, only to continue to sweat the next day.since then (it's been about four days) i have applied myself every night without any decrease in the amount i sweat.today was the first day of school and i was sweat the entire day, unable to lift my $F(\boldsymbol{s}_{\text{adv}}, l =$"2.0") = 0.55 **Cos.** = 0.01 $MedSTS$ = 1.0	'i started using this mask a little more than a week ago.i applied it three nights in a row as instructed, and went to a clinic the next day to test it out.i still sweated, but not nearly as much, and i kept convinced that with time i would be totally sweat free.i applied it once again the following night, only to continue to sweat the next day.since then (it's been about four days) i have applied hypercare every night without any changes in the amount i sweat.today was the first day of school and i was sweat the entire day, unable to lift my $F(\boldsymbol{s}_{\text{adv}}, l =$"2.0") = 1.0 **Cos.** = 0.42 $MedSTS$ = 0.91
Original	'not every medicine is for everyone, but as one who has tried most of the major pharmaceuticals for major depression, panic attacks, severe anxiety and anxiety related bouts of obsessive compulsive disorder, i can tell you lexapro is the only medicine that i've been able to stay on and be effective for my mental well-being...it is the only one i've had no side effects with. other ssri's have either: made me more anxious and/or depressed, dry mouth, bad weight gain, or extreme fatigue making me into a walking zombie during the day. i've been on lexapro 6 years $F(\boldsymbol{s}, l =$"10.0") = 1.0	'not every medicine is for everyone, but as one who has tried most of the major pharmaceuticals for major depression, panic attacks, severe anxiety and anxiety related bouts of obsessive compulsive disorder, i can tell you lexapro is the only medicine that i've been able to stay on and be effective for my mental well-being...it is the only one i've had no side effects with. other ssri's have either: made me more anxious and/or depressed, dry mouth, bad weight gain, or extreme fatigue making me into a walking zombie during the day. i've been on lexapro 6 years $F(\boldsymbol{s}, l =$"10.0") = 0.99	'not every medicine is for everyone, but as one who has tried most of the major pharmaceuticals for major depression, panic attacks, severe anxiety and anxiety related bouts of obsessive compulsive disorder, i can tell you lexapro is the only medicine that i've been able to stay on and be effective for my mental well-being...it is the only one i've had no side effects with. other ssri's have either: made me more anxious and/or depressed, dry mouth, bad weight gain, or extreme fatigue making me into a walking zombie during the day. i've been on lexapro 6 years $F(\boldsymbol{s}, l =$"10.0") = 1.0

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	'not every medicine is for everyone, but as one who has tried most of the major pharmaceuticals for major depression, panic attacks, severe anxiety and anxiety related bouts of obsessive compulsive disorder, i can reassure you lexapro is the only medicine that i've been able to stay on and be effective for my personal well-being...it is the only one i've had no side effects with. My ssri's have never made me more anxious and/or depressed, dry mouth, bad weight gain, or extreme fatigue making me into a walking zombie during the day. i've been on lexapro 6 years $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"10.0"}) = 1.0$ **Cos.** = -0.2 $MedSTS = 0.97$	'not every medicine is for everyone, but as one who has prescribed most of the major pharmaceuticals for major depression, panic attacks, severe anxiety and anxiety related bouts of obsessive compulsive disorder, i can tell you lexapro is the only medicine that i've been able to stay on and be effective for my mental well-being...it is the only one i've had no side effects . pill antidepressants have either: made me more anxious and/or depressed, dry mouth, bad weight gain, or extreme fatigue making me into a walking zombie during the day. i've been on lexapro 6 years $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"10.0"}) = 1.0$ **Cos.** = 0.04 $MedSTS = 0.95$	'not every medicine is for everyone, but as one who has done most of the major medication for major depression, panic attacks, severe anxiety and anxiety related bouts of obsessive compulsive disorder, i can tell you lexapro is the only medicine that i've been able to focus on and be positive for my mental well-being...it is the only one i've had no side effects with. other ssri's have either: made me more anxious and/or depressed, dry mouth, bad weight gain, or extreme fatigue making me into a walking zombie during the day. i've been on lexapro 6 years $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"10.0"}) = 1.0$ **Cos.** = 0.52 $MedSTS = 0.92$
Original	'just took my first dose 5 mg of brintellix - have been on every possible medication including wellbutrin for 15 years, seroquel for 9 years, lexapro for 2 years, just weaned off lexapro.i feel quite odd, butterflies in stomach and brain fog - my daughter has been on brintellix for 2 months and is still vomiting - if this continues, another failed med.' $F(\boldsymbol{s},\ l=\text{"4.0"}) = 1.0$	'just took my first dose 5 mg of brintellix - have been on every possible medication including wellbutrin for 15 years, seroquel for 9 years, lexapro for 2 years, just weaned off lexapro.i feel quite odd, butterflies in stomach and brain fog - my daughter has been on brintellix for 2 months and is still vomiting - if this continues, another failed med.' $F(\boldsymbol{s},\ l=\text{"4.0"}) = 1.0$	'just took my first dose 5 mg of brintellix - have been on every possible medication including wellbutrin for 15 years, seroquel for 9 years, lexapro for 2 years, just weaned off lexapro.i feel quite odd, butterflies in stomach and brain fog - my daughter has been on brintellix for 2 months and is still vomiting - if this continues, another failed med.' $F(\boldsymbol{s},\ l=\text{"4.0"}) = 1.0$
Adversarial	'just took my first full 5 mg of brintellix - have been on every possible medication including wellbutrin for 15 years, seroquel for 9 years, lexapro for 2 years, just weaned off lexapro.i feel quite and butterflies in stomach and brain fog - my daughter has been on brintellix for 2 months and is still vomiting - if this continues, another new med.' $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"4.0"}) = 0.6$ **Cos.** = -0.18 $MedSTS = 1.0$	'just took my first batch 5 mg of brintellix - have been on every possible medication including wellbutrin for 15 years, seroquel for 9 years, lexapro for 2 years, just weaned off lexapro.i feel quite sick butterflies in stomach and brain fog - my daughter has been on brintellix for 2 months and is still vomiting - if this continues, another miracle med.' $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"4.0"}) = 0.76$ **Cos.** = 0.04 $MedSTS = 1.0$	'just took my first full 5 mg of brintellix - have been on every possible medication including wellbutrin for 15 years, seroquel for 9 years, lexapro for 2 years, just weaned off lexapro.i feel quite odd, butterflies in stomach and brain fog - my daughter has been on brintellix for 2 months and is still awake - if this continues, another antidepressant med.' $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"4.0"}) = 1.0$ **Cos.** = 0.5 $MedSTS = 1.0$
Original	'after trying zoloft and lexapro, without any success and made my symptoms worse. luvox helped me getting my life back, the best medicine. i feel much more in control of my ocd. excellent.. even when i feel sleepy sometimes as a side effect.. its worth it!' $F(\boldsymbol{s},\ l=\text{"10.0"}) = 1.0$	'after trying zoloft and lexapro, without any success and made my symptoms worse. luvox helped me getting my life back, the best medicine. i feel much more in control of my ocd. excellent.. even when i feel sleepy sometimes as a side effect.. its worth it!' $F(\boldsymbol{s},\ l=\text{"10.0"}) = 1.0$	'after trying zoloft and lexapro, without any success and made my symptoms worse. luvox helped me getting my life back, the best medicine. i feel much more in control of my ocd. excellent.. even when i feel sleepy sometimes as a side effect.. its worth it!' $F(\boldsymbol{s},\ l=\text{"10.0"}) = 1.0$
Adversarial	'after trying zoloft and lexapro, without any success and made my symptoms worse. luvox helped me getting my life back, the best medicine. i feel much more in spite of my euph excellent.. even when i feel sleepy sometimes as a side effect.. its worth it!' $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"10.0"}) = 1.0$ **Cos.** = -0.34 $MedSTS = 0.79$	'after trying zoloft and lexapro, without any success and made my symptoms worse. luvox helped me getting my life back, the best medicine. i feel much more in spite of my sleeping excellent.. even when i feel sleepy sometimes as a side effect.. its worth it!' $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"10.0"}) = 1.0$ **Cos.** = -0.08 $MedSTS = 0.81$	'after trying zoloft and lexapro, without any success and made my symptoms worse. luvox helped me getting my life back, the best thing i feel much more in control of my ocd. excellent.. even when i feel pain sometimes as a side effect.. its worth it!' $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"10.0"}) = 1.0$ **Cos.** = 0.41 $MedSTS = 0.8$
Original	'i was put on tri sprintec when i started getting my periods every two weeks, and was on it for three months. it fixed my period problem, but i had never had a problem with acne until starting this. my acne got so much worse and would clear up instantly once i started the sugar pills. i went from a d to a dd which is kind of annoying but i didn't gain much weight at least. the worst part, however, was the tenderness in my breasts, it was horrible. painful to the touch, running or going to the gym was horribly uncomfortable. just like the acne, during the $F(\boldsymbol{s},\ l=\text{"6.0"}) = 1.0$	'i was put on tri sprintec when i started getting my periods every two weeks, and was on it for three months. it fixed my period problem, but i had never had a problem with acne until starting this. my acne got so much worse and would clear up instantly once i started the sugar pills. i went from a d to a dd which is kind of annoying but i didn't gain much weight at least. the worst part, however, was the tenderness in my breasts, it was horrible. painful to the touch, running or going to the gym was horribly uncomfortable. just like the acne, during the $F(\boldsymbol{s},\ l=\text{"6.0"}) = 0.92$	'i was put on tri sprintec when i started getting my periods every two weeks, and was on it for three months. it fixed my period problem, but i had never had a problem with acne until starting this. my acne got so much worse and would clear up instantly once i started the sugar pills. i went from a d to a dd which is kind of annoying but i didn't gain much weight at least. the worst part, however, was the tenderness in my breasts, it was horrible. painful to the touch, running or going to the gym was horribly uncomfortable. just like the acne, during the $F(\boldsymbol{s},\ l=\text{"6.0"}) = 1.0$

	Vanilla	Adversarial	FAR-IG
Adversarial	'i was put on tri sprintec when i started getting my periods every two weeks, and was on it for three months. it cured my period problem, but i had never had a problem with sugar until starting this. my stomach got so much better and would clear up instantly once i started the sugar pills. i went from a d to a dd which is kind of annoying but i didn't gain much weight at first the worst part, however, was the tenderness in my breasts, it was horrible. painful to the touch, running or going to the gym was horribly uncomfortable. just like the acne, during the $F(\boldsymbol{s}_{\text{adv}},\ l=\text{“6.0”}) = 0.98$ **Cos.** = -0.37 $MedSTS = 0.76$	'i was put on tri sprintec when i started getting my periods every two weeks, and was on it for three months. it caused my period problem, but i had never had a problem with this until starting this. my stomach got so much easier and would clear up immediately once i started the sugar pills. i went from a d to a dd which is kind of annoying but i didn't gain much weight at least. the worst part, however, was the tenderness in my breasts, it was horrible. painful to the touch, running or going to the gym was horribly uncomfortable. just like the acne, during the $F(\boldsymbol{s}_{\text{adv}},\ l=\text{“6.0”}) = 0.72$ **Cos.** = -0.11 $MedSTS = 0.75$	'i was put on tri sprintec when i started getting my periods every two weeks, and was on it for three months. it fixed my period problem, but i had never had a problem with this until starting this. my headaches got so much worse and would clear up instantly once i started the sugar syrup which ranging from a d to a dd which is kind of annoying but i didn't gain much weight at least. the worst part, however, was the tenderness in my breasts, it was horrible. painful to the touch, running or going to the gym was horribly uncomfortable. just like the acne, during the $F(\boldsymbol{s}_{\text{adv}},\ l=\text{“6.0”}) = 0.97$ **Cos.** = 0.68 $MedSTS = 0.82$
Original	'mestinon helps everyone differently. i started out with ocular mg. apparently, this drug helps most people with their ocular mg, but it doesn't do anything at all to improve my eye. my case of mg rapidly generalized, and difficulty breathing was my 2nd symptom to manifest. mestinon improves my breathing issues somewhat, but doesn't take the shortness of breath away completely. same with my arms and thighs; it helps, but doesn't make the weakness disappear altogether. mestinon can cause diarrhea, but most likely won't if taken alongside a meal or with a small snack. $F(\boldsymbol{s},\ l=\text{“6.0”}) = 1.0$	'mestinon helps everyone differently. i started out with ocular mg. apparently, this drug helps most people with their ocular mg, but it doesn't do anything at all to improve my eye. my case of mg rapidly generalized, and difficulty breathing was my 2nd symptom to manifest. mestinon improves my breathing issues somewhat, but doesn't take the shortness of breath away completely. same with my arms and thighs; it helps, but doesn't make the weakness disappear altogether. mestinon can cause diarrhea, but most likely won't if taken alongside a meal or with a small snack. $F(\boldsymbol{s},\ l=\text{“6.0”}) = 1.0$	'mestinon helps everyone differently. i started out with ocular mg. apparently, this drug helps most people with their ocular mg, but it doesn't do anything at all to improve my eye. my case of mg rapidly generalized, and difficulty breathing was my 2nd symptom to manifest. mestinon improves my breathing issues somewhat, but doesn't take the shortness of breath away completely. same with my arms and thighs; it helps, but doesn't make the weakness disappear altogether. mestinon can cause diarrhea, but most likely won't if taken alongside a meal or with a small snack. $F(\boldsymbol{s},\ l=\text{“6.0”}) = 1.0$
Adversarial	'mestinon helps everyone .' i started out with ocular mg. apparently, this drug helps most kids with their ocular mg, but it doesn't do anything at all to improve my eye. my case of mg is generalized, and difficulty breathing was my 2nd symptom to manifest. mestinon improves my breathing issues somewhat, but doesn't take the shortness of breath away completely. same with my arms and thighs; it helps, but doesn't make the nausea disappear altogether. mestinon can cause diarrhea, but most likely won't if taken alongside a meal or with a small snack. $F(\boldsymbol{s}_{\text{adv}},\ l=\text{“6.0”}) = 1.0$ **Cos.** = -0.08 $MedSTS = 0.92$	'mestinon helps everyone ' i started out with ocular mg. apparently, this supplement helps most people with their ocular mg, but it doesn't do anything at all to relieve my eye. my intake of mg rapidly generalized, and difficulty breathing was my 2nd symptom to manifest. mestinon improves my breathing issues somewhat, but doesn't take the shortness of breath away completely. same with my arms and thighs; it helps, but doesn't make the weakness disappear altogether. mestinon can cause diarrhea, but most likely won't if taken alongside a meal or with a small snack. $F(\boldsymbol{s}_{\text{adv}},\ l=\text{“6.0”}) = 0.99$ **Cos.** = 0.38 $MedSTS = 0.94$	'mestinon helps everyone .' i started out with ocular mg. apparently, this pill helps most people with their ocular mg, but it doesn't do anything at all to improve my eye. my case of mg rapidly generalized, and difficulty breathing was my 2nd cause to manifest. mestinon improves my breathing issues somewhat, but doesn't take the shortness of coughing away completely. same with my arms and thighs; it helps, but doesn't make the weakness disappear altogether. mestinon can cause diarrhea, but most likely won't if taken alongside a meal or with a small snack. $F(\boldsymbol{s}_{\text{adv}},\ l=\text{“6.0”}) = 1.0$ **Cos.** = 0.7 $MedSTS = 0.95$
Original	'picked up a nasty h pylori strain from an casual blind date, i know i should have gotten to know the person better. took a while before symptoms showed up. had severe upset stomach, occasional diarrhea, nausea and slow but steady weight loss. took a long time and several doctors to diagnose my steadily worsening condition. tried prevpak first, seemed to work at first butmy infection came back. the new gi then prescribed pylera after my 3rd endoscopy. pylera has worked, it's been a year and i am still h pylera negative. but it's been brutal and $F(\boldsymbol{s},\ l=\text{“6.0”}) = 1.0$	'picked up a nasty h pylori strain from an casual blind date, i know i should have gotten to know the person better. took a while before symptoms showed up. had severe upset stomach, occasional diarrhea, nausea and slow but steady weight loss. took a long time and several doctors to diagnose my steadily worsening condition. tried prevpak first, seemed to work at first butmy infection came back. the new gi then prescribed pylera after my 3rd endoscopy. pylera has worked, it's been a year and i am still h pylera negative. but it's been brutal and $F(\boldsymbol{s},\ l=\text{“6.0”}) = 0.98$	'picked up a nasty h pylori strain from an casual blind date, i know i should have gotten to know the person better. took a while before symptoms showed up. had severe upset stomach, occasional diarrhea, nausea and slow but steady weight loss. took a long time and several doctors to diagnose my steadily worsening condition. tried prevpak first, seemed to work at first butmy infection came back. the new gi then prescribed pylera after my 3rd endoscopy. pylera has worked, it's been a year and i am still h pylera negative. but it's been brutal and $F(\boldsymbol{s},\ l=\text{“6.0”}) = 0.9$

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	'picked up a nasty h pylori ' from an casual doctor date, i know i should have gotten to know the person better. took a while until they showed up. had severe upset stomach, occasional diarrhea, nausea and slow but steady weight loss. took a long time and several doctors to diagnose my steadily worsening condition. tried prevpak first, seemed to work at first butmy infection came back. the new gi then prescribed pylera after my 3rd endoscopy. pylera has worked, it's been a year and i am still h pylera negative. but it's been brutal and $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"6.0"}) = 0.97$ **Cos.** = -0.19 $MedSTS = 0.98$	'picked up a nasty h pylori pill from an anonymous internet date, i know i should have gotten to know the person better. took a while until symptoms showed up. had severe upset stomach, occasional diarrhea, nausea and slow but steady weight loss. took a long time and several doctors to diagnose my steadily worsening condition. tried prevpak first, seemed to work at first butmy infection came back. the new gi then prescribed pylera after my 3rd endoscopy. pylera has worked, it's been a year and i am still h pylera negative. but it's been brutal and $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"6.0"}) = 0.82$ **Cos.** = 0.19 $MedSTS = 0.9$	'picked up a nasty h pylori rash from an infected infection so i know i should have gotten to know the person better. took a while before symptoms showed up. had severe upset stomach, occasional diarrhea, nausea and slow but steady weight loss. took a long time and several doctors to diagnose my steadily worsening condition. tried prevpak first, seemed to work at first butmy infection came back. the new gi then prescribed pylera after my 3rd endoscopy. pylera has worked, it's been a year and i am still h pylera negative. but it's been brutal and $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"6.0"}) = 0.78$ **Cos.** = 0.52 $MedSTS = 0.81$
Original	'mom is 98 years old, the gabapentin capsules was for her knees pain, 100 mg, twice a day. it did help for her pain relief, however we foundshe started having bed side effect as dark urine, feeling tire, not hungry, feeling confused at night time, so we decided to stop the medication.' $F(\boldsymbol{s},\ l=\text{"4.0"}) = 1.0$	'mom is 98 years old, the gabapentin capsules was for her knees pain, 100 mg, twice a day. it did help for her pain relief, however we foundshe started having bed side effect as dark urine, feeling tire, not hungry, feeling confused at night time, so we decided to stop the medication.' $F(\boldsymbol{s},\ l=\text{"4.0"}) = 1.0$	'mom is 98 years old, the gabapentin capsules was for her knees pain, 100 mg, twice a day. it did help for her pain relief, however we foundshe started having bed side effect as dark urine, feeling tire, not hungry, feeling confused at night time, so we decided to stop the medication.' $F(\boldsymbol{s},\ l=\text{"4.0"}) = 1.0$
Adversarial	'mom is 8 years old, the gabapentin capsules was for her knees pain, 100 mg, twice a day. it did help for her pain relief, however we foundshe started having bed side effect as dark urine, feeling tire, not hungry, feeling confused at night time, so we decided to try the medication.' $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"4.0"}) = 1.0$ **Cos.** = 0.23 $MedSTS = 0.76$	'mom is 5 years old, the gabapentin capsules was for her knees pain, 100 mg, twice a day. it did help for her pain relief, however we foundshe started having bed side effect as dark nights feeling tire, not hungry, feeling confused at night time, so we decided to stop the medication.' $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"4.0"}) = 1.0$ **Cos.** = 0.44 $MedSTS = 0.77$	'mom about 16 years old, the gabapentin capsules was for her knees pain, 100 mg, twice a day. it did help for her pain relief, however we foundshe started having bed side effect as dark urine, feeling tire, not hungry, feeling confused at night time, so we decided to stop the medication.' $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"4.0"}) = 1.0$ **Cos.** = 0.65 $MedSTS = 0.76$
Original	'i was hit by a police car in 2002. i was in recovery for a year. from the head down. broken jaw, 4 disks crushed my neck. both collar bones, 7 broken ribs and a shatered pelvis. so many pills, patches and injections. under the knife many times. a real humpty dumpty story. i have more pins, screws in me than your local hard wear store. note: i don't want to over state the how effective it is but it works. talking to my doctor about the pain. he gave me this to try. well my 1st thought was &qu $F(\boldsymbol{s},\ l=\text{"8.0"}) = 1.0$	'i was hit by a police car in 2002. i was in recovery for a year. from the head down. broken jaw, 4 disks crushed my neck. both collar bones, 7 broken ribs and a shatered pelvis. so many pills, patches and injections. under the knife many times. a real humpty dumpty story. i have more pins, screws in me than your local hard wear store. note: i don't want to over state the how effective it is but it works. talking to my doctor about the pain. he gave me this to try. well my 1st thought was &qu $F(\boldsymbol{s},\ l=\text{"8.0"}) = 0.95$	'i was hit by a police car in 2002. i was in recovery for a year. from the head down. broken jaw, 4 disks crushed my neck. both collar bones, 7 broken ribs and a shatered pelvis. so many pills, patches and injections. under the knife many times. a real humpty dumpty story. i have more pins, screws in me than your local hard wear store. note: i don't want to over state the how effective it is but it works. talking to my doctor about the pain. he gave me this to try. well my 1st thought was &qu $F(\boldsymbol{s},\ l=\text{"8.0"}) = 0.92$
Adversarial	'i was hit by a police car in 2002. i was in jail for a year. from the head down. broken jaw, 4 disks crushed my neck. both collar bones, 7 broken ribs and a shatered pelvis. so many pills, patches and injections. under the covers many times. a real humpty dumpty story. i have more pins, screws in me than your local hard wear store. note: i don't want to over compl the how horrible it is but it works. lied to my doctor about the pain. he gave me this to try. well my 1st thought was &qu $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"8.0"}) = 1.0$ **Cos.** = -0.35 $MedSTS = 0.9$	'i was stopped by a police car in 2002. i was in rehab for a year. from the head down. broken jaw, 4 disks crushed my neck. both collar bones, 7 broken ribs and a shatered pelvis. so many pills, patches and injections. under the counter many times. a real humpty dumpty story. i have more pins, needles in me than your local hard wear store. note: i don't want to over state the how horrible it is but it works. talking to my doctor about the pain. he gave me this to try. well my 1st thought was &qu $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"8.0"}) = 0.7$ **Cos.** = 0.07 $MedSTS = 0.84$	'i was diagnosed by a chirop surgeon in 2002. i was in surgery for a year. from the head down. broken jaw, 4 disks crushed my neck. both collar bones, 7 broken ribs and a shatered pelvis. so many pills, patches and injections. under the knife many times. a real humpty dumpty story. i have more pins, stuck in me than your local hard wear store. note: i don't want to over state the how effective it is but it works. talking to my doctor about the pain. he gave me this to try. well my 1st thought was &qu $F(\boldsymbol{s}_{\text{adv}},\ l=\text{"8.0"}) = 0.52$ **Cos.** = 0.65 $MedSTS = 0.69$

	VANILLA	ADVERSARIAL	FAR-IG
Original	'been on antidepressants since around 1991 after i had thyroid removed.taking wellbutrin for depression which works okay but last year started having late night anxiety attacks.really bad fear.i don't understand why.had me walking the streets at 11pm.started viibryd and had weight gain and stomach issues and by the time i got to 40mg the anxiety just stopped! this medicine does as advertised for me, it stops the anxiety. that's what i needed and that's what i got. like a miracle.' $F(\boldsymbol{s}, l=\text{“10.0”}) = 1.0$	'been on antidepressants since around 1991 after i had thyroid removed.taking wellbutrin for depression which works okay but last year started having late night anxiety attacks.really bad fear.i don't understand why.had me walking the streets at 11pm.started viibryd and had weight gain and stomach issues and by the time i got to 40mg the anxiety just stopped! this medicine does as advertised for me, it stops the anxiety. that's what i needed and that's what i got. like a miracle.' $F(\boldsymbol{s}, l=\text{“10.0”}) = 1.0$	'been on antidepressants since around 1991 after i had thyroid removed.taking wellbutrin for depression which works okay but last year started having late night anxiety attacks.really bad fear.i don't understand why.had me walking the streets at 11pm.started viibryd and had weight gain and stomach issues and by the time i got to 40mg the anxiety just stopped! this medicine does as advertised for me, it stops the anxiety. that's what i needed and that's what i got. like a miracle.' $F(\boldsymbol{s}, l=\text{“10.0”}) = 1.0$
Adversarial	'been on lithium since around 1991 after i had thyroid removed.taking wellbutrin for depression which works okay but last year started having late night anxiety attacks.really bad fear.i don't understand why.had me walking the streets at 11pm.started viibryd and had weight gain and stomach issues and by the time i got to 40mg the anxiety just stopped! this medicine does as nothing for me, it stops the anxiety. that's what i needed and that's what i got. have a reply $F(\boldsymbol{s}_{\text{adv}}, l=\text{“10.0”}) = 0.98$ **Cos.** = 0.05 $MedSTS = 0.91$	'been on antidepressants since around 2012 after i had thyroid removed.taking wellbutrin for depression which works okay but last year started having late night heart attacks.really bad fear.i don't understand why.had me walking the streets at 11pm.started viibryd and had weight gain and stomach issues and by the time i got to eat the pain just stopped! this medicine does as advertised for me, it stops the anxiety. that's what i needed and that's what i got. like a miracle.' $F(\boldsymbol{s}_{\text{adv}}, l=\text{“10.0”}) = 1.0$ **Cos.** = 0.53 $MedSTS = 0.8$	'been on lithium since around 1991 after i had thyroid removed.taking wellbutrin for arthritis which works okay but last year started having late night sleep attacks.really bad fear.i don't understand why.had me walking the streets at 11pm.started viibryd and had weight gain and stomach issues and by the time i got to 40mg the symptoms just stopped! this medicine does as advertised for me, it stops the anxiety. that's what i needed and that's what i got. like a miracle.' $F(\boldsymbol{s}_{\text{adv}}, l=\text{“10.0”}) = 0.96$ **Cos.** = 0.77 $MedSTS = 0.78$
Original	'i'm a nurse and know all about package inserts explaining side effects but man they weren't kidding. i was started about 3 weeks ago by my doctor who said i needed to lose a bunch of weight. i expected some stomach upset but not like this. now i understand why the slow titration to higher more effective dosing. i can't bear the side effects. i cannot work with stomach upset all day and need to rest at night. back to diet and exercise for me. this is just not right.' $F(\boldsymbol{s}, l=\text{“2.0”}) = 1.0$	'i'm a nurse and know all about package inserts explaining side effects but man they weren't kidding. i was started about 3 weeks ago by my doctor who said i needed to lose a bunch of weight. i expected some stomach upset but not like this. now i understand why the slow titration to higher more effective dosing. i can't bear the side effects. i cannot work with stomach upset all day and need to rest at night. back to diet and exercise for me. this is just not right.' $F(\boldsymbol{s}, l=\text{“2.0”}) = 1.0$	'i'm a nurse and know all about package inserts explaining side effects but man they weren't kidding. i was started about 3 weeks ago by my doctor who said i needed to lose a bunch of weight. i expected some stomach upset but not like this. now i understand why the slow titration to higher more effective dosing. i can't bear the side effects. i cannot work with stomach upset all day and need to rest at night. back to diet and exercise for me. this is just not right.' $F(\boldsymbol{s}, l=\text{“2.0”}) = 0.99$
Adversarial	'i'm a nurse and know all about package inserts explaining side effects but man they weren't helpful i was started about 3 weeks ago by my doctor who said i needed to lose a bunch of weight. i expected some stomach discomfort but not like this. now i understand why the slow titration to higher more effective dosing. i can't explain the side effects. i cannot work with stomach upset all day and need to rest at night. back to diet and exercise for me. this is just totally right.' $F(\boldsymbol{s}_{\text{adv}}, l=\text{“2.0”}) = 1.0$ **Cos.** = -0.02 $MedSTS = 0.88$	'i'm a nurse and know all about package inserts explaining side effects but man they weren't helpful i was started about 3 weeks ago by my doctor who said i needed to lose a bunch of weight. i expected some stomach problems but not like this. now i understand why the slow titration to higher more insulin dosing. i can't explain the side effects. i cannot work with stomach upset all day and need to rest at night. back to diet and exercise for me. this is just not right.' $F(\boldsymbol{s}_{\text{adv}}, l=\text{“2.0”}) = 0.95$ **Cos.** = 0.28 $MedSTS = 0.87$	'i'm a nurse and know all about package pills explaining side effects but man they weren't kidding. i was started about 3 weeks ago by my doctor who said i needed to lose a bunch of weight. i expected some stomach problems but not like this. now i understand why the slow titration to higher more effective dosing. i can't bear the side effects. i cannot work with stomach sickness all day and need to rest at night. back to bed and exercise for me. this is just not right.' $F(\boldsymbol{s}_{\text{adv}}, l=\text{“2.0”}) = 0.99$ **Cos.** = 0.63 $MedSTS = 0.93$
Original	'my husband got this for me since i haven't been pooping that well lately. he told me it was better than any others, because you don't get stomach cramps. wrong. i took 2 a couple of hours ago, and now, my stomach is in the worst pain ever. thankfully, i have had a bm, but i don't think i will ever take dulcolax again because of the severe pain.' $F(\boldsymbol{s}, l=\text{“4.0”}) = 1.0$	'my husband got this for me since i haven't been pooping that well lately. he told me it was better than any others, because you don't get stomach cramps. wrong. i took 2 a couple of hours ago, and now, my stomach is in the worst pain ever. thankfully, i have had a bm, but i don't think i will ever take dulcolax again because of the severe pain.' $F(\boldsymbol{s}, l=\text{“4.0”}) = 1.0$	'my husband got this for me since i haven't been pooping that well lately. he told me it was better than any others, because you don't get stomach cramps. wrong. i took 2 a couple of hours ago, and now, my stomach is in the worst pain ever. thankfully, i have had a bm, but i don't think i will ever take dulcolax again because of the severe pain.' $F(\boldsymbol{s}, l=\text{“4.0”}) = 1.0$

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	'my husband got this for me since i haven't been pooping that well lately. he told me it was better than any others, because you don't get stomach cramps. wrong. i took 2 a couple of hours ago, and now, my stomach is in the worst pain ever. Recently i have had a bm, but i don't think i will wanna take dulcolax again because of the constant pain.' $F(\boldsymbol{s}_{\mathrm{adv}},\ l=\text{“4.0”}) = 1.0$ **Cos.** = 0.16 $MedSTS = 0.98$	'my husband got this for me since i haven't been pooping that well lately. he told me it was better than any others, because you don't get stomach cramps. wrong. i took 2 a couple of hours ago, and now, my stomach is in the worst pain ever. Recently i have had a relapse but i don't think i will ever take dulcolax again because of the constant pain.' $F(\boldsymbol{s}_{\mathrm{adv}},\ l=\text{“4.0”}) = 1.0$ **Cos.** = 0.39 $MedSTS = 0.98$	'my surgeon got this for me since i haven't been exercising that well lately. he told me it was better than any others, because you don't get stomach cramps. wrong. i took 2 a couple of hours ago, and now, my butt is in the worst pain ever. thankfully, i have had a bm, but i don't think i will ever take dulcolax again because of the severe pain.' $F(\boldsymbol{s}_{\mathrm{adv}},\ l=\text{“4.0”}) = 0.99$ **Cos.** = 0.69 $MedSTS = 0.73$
Original	'my husband takes saphris 5 mg twice a day for bipolar depression. he was having delusions and psychosis as well as paranoia. since starting the saphris these have stopped, but there are times when he is very stoned acting, like a zombie. he isn't responsive or talkative a lot of the time. his sex drive has also become non existent almost and he suffers from ed when we do have sex.it's so difficult when you have a disease like bd and then you have to take these meds, which help some of the symptoms but bring on others that are unwanted. it's like $F(\boldsymbol{s},\ l=\text{“8.0”}) = 1.0$	'my husband takes saphris 5 mg twice a day for bipolar depression. he was having delusions and psychosis as well as paranoia. since starting the saphris these have stopped, but there are times when he is very stoned acting, like a zombie. he isn't responsive or talkative a lot of the time. his sex drive has also become non existent almost and he suffers from ed when we do have sex.it's so difficult when you have a disease like bd and then you have to take these meds, which help some of the symptoms but bring on others that are unwanted. it's like $F(\boldsymbol{s},\ l=\text{“8.0”}) = 0.66$	'my husband takes saphris 5 mg twice a day for bipolar depression. he was having delusions and psychosis as well as paranoia. since starting the saphris these have stopped, but there are times when he is very stoned acting, like a zombie. he isn't responsive or talkative a lot of the time. his sex drive has also become non existent almost and he suffers from ed when we do have sex.it's so difficult when you have a disease like bd and then you have to take these meds, which help some of the symptoms but bring on others that are unwanted. it's like $F(\boldsymbol{s},\ l=\text{“8.0”}) = 0.88$
Adversarial	'my husband takes saphris 5 mg twice a day for clinical depression. he was having delusions and psychosis as well as paranoia. since starting the saphris these have stopped, but there are times when he is very strangely acting, like a zombie. he isn't lucid or talkative a lot of the time his sex drive has also become non existent almost and he suffers from ed when we do have sex.it's so frustrating when you have a disease like bd and then you have to take these meds, which help some of the symptoms but bring on others that are unwanted. it's like $F(\boldsymbol{s}_{\mathrm{adv}},\ l=\text{“8.0”}) = 0.98$ **Cos.** = -0.31 $MedSTS = 0.91$	'my self takes saphris 5 mg twice a day for clinical depression. he was having delusions and psychosis as well as paranoia. since starting the saphris these have stopped, but there are times when he is very strange acting, like a zombie. he isn't lucid or talkative a lot of the time. his sex disorder has also become non existent almost and he suffers from ed when we do have sex.it's so difficult when you have a disease like bd and then you have to take these meds, which help some of the symptoms but bring on others that are unwanted. it's like $F(\boldsymbol{s}_{\mathrm{adv}},\ l=\text{“8.0”}) = 0.55$ **Cos.** = -0.08 $MedSTS = 0.88$	'my self takes saphris 5 x twice a day for mood depression. he was having delusions and psychosis as well as paranoia. since starting the saphris these have stopped, but there are times when he is very stoned acting, like a zombie. he isn't responsive or talkative a lot of the time. his manic disorder has also become non existent almost and he suffers from ed when we do have sex.it's so difficult when you have a disease like bd and then you have to take these meds, which help some of the symptoms but bring on others that are unwanted. it's like $F(\boldsymbol{s}_{\mathrm{adv}},\ l=\text{“8.0”}) = 0.87$ **Cos.** = 0.47 $MedSTS = 0.88$
Original	background thrombospondin - 1 (tsp - 1) promotes breast cancer cell invasion of collagen by upregulating matrix metalloproteinase - 9 (mmp - 9) production . stromal tsp - 1 may play a role in regulating tumor cell invasion . we hypothesize that fibroblasts promote breast cancer cell invasion by upregulating the production of mmp - 9 through tsp - 1 . methods mda - mb - 231 human breast carcinoma cells were grown alone or in coculture with human fibroblasts . gelatin zymography and western immunoblot analysis for mmp - 9 were performed on the coculture cell media and the single cell media . inhibition of fibroblast - mediated breast tumor cell invasion by an anti - tsp - 1 or an anti - mmp - 9 antibody was evaluated using a modified boyden chamber . results coculture experiments showed an increased production of mmp - 9 when compared with breast cancer single cell culture or fibroblast single cell culture experiments as demonstrated by zymography and western immunoblot analysis . fibroblast - stimulated mmp - 9 production was comparable with tsp - 1 - stimulated mmp - 9 production . anti - tsp - 1 antibody and anti - mmp - 9 antibody inhibited fibroblast - stimulated tumor cell invasion to 30 % and 26 % of controls , respectively (p < . 05) . conclusions fibroblasts may regulate breast cancer cell invasion by promoting tumor mmp - 9 production $F(\boldsymbol{s},\ l=$“<multilabel>”$) = 1.0$	background thrombospondin - 1 (tsp - 1) promotes breast cancer cell invasion of collagen by upregulating matrix metalloproteinase - 9 (mmp - 9) production . stromal tsp - 1 may play a role in regulating tumor cell invasion . we hypothesize that fibroblasts promote breast cancer cell invasion by upregulating the production of mmp - 9 through tsp - 1 . methods mda - mb - 231 human breast carcinoma cells were grown alone or in coculture with human fibroblasts . gelatin zymography and western immunoblot analysis for mmp - 9 were performed on the coculture cell media and the single cell media . inhibition of fibroblast - mediated breast tumor cell invasion by an anti - tsp - 1 or an anti - mmp - 9 antibody was evaluated using a modified boyden chamber . results coculture experiments showed an increased production of mmp - 9 when compared with breast cancer single cell culture or fibroblast single cell culture experiments as demonstrated by zymography and western immunoblot analysis . fibroblast - stimulated mmp - 9 production was comparable with tsp - 1 - stimulated mmp - 9 production . anti - tsp - 1 antibody and anti - mmp - 9 antibody inhibited fibroblast - stimulated tumor cell invasion to 30 % and 26 % of controls , respectively (p < . 05) . conclusions fibroblasts may regulate breast cancer cell invasion by promoting tumor mmp - 9 production $F(\boldsymbol{s},\ l=$“<multilabel>”$) = 1.0$	background thrombospondin - 1 (tsp - 1) promotes breast cancer cell invasion of collagen by upregulating matrix metalloproteinase - 9 (mmp - 9) production . stromal tsp - 1 may play a role in regulating tumor cell invasion . we hypothesize that fibroblasts promote breast cancer cell invasion by upregulating the production of mmp - 9 through tsp - 1 . methods mda - mb - 231 human breast carcinoma cells were grown alone or in coculture with human fibroblasts . gelatin zymography and western immunoblot analysis for mmp - 9 were performed on the coculture cell media and the single cell media . inhibition of fibroblast - mediated breast tumor cell invasion by an anti - tsp - 1 or an anti - mmp - 9 antibody was evaluated using a modified boyden chamber . results coculture experiments showed an increased production of mmp - 9 when compared with breast cancer single cell culture or fibroblast single cell culture experiments as demonstrated by zymography and western immunoblot analysis . fibroblast - stimulated mmp - 9 production was comparable with tsp - 1 - stimulated mmp - 9 production . anti - tsp - 1 antibody and anti - mmp - 9 antibody inhibited fibroblast - stimulated tumor cell invasion to 30 % and 26 % of controls , respectively (p < . 05) . conclusions fibroblasts may regulate breast cancer cell invasion by promoting tumor mmp - 9 production $F(\boldsymbol{s},\ l=$“<multilabel>”$) = 1.0$

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	introduction thrombospondin - 1 (tsp - 1) promotes breast cancer cell invasion of collagen by upregulating matrix metalloproteinase - 9 (mmp - 9) production . tumor tsp - 1 may play a role in regulating tumor cell growth . we hypothesize that they promote breast cancer cell proliferation by upregulating the production of mmp - 9 through mmp - 1 . methods mda - mb - 231 human breast carcinoma cells were grown alone or in coculture with human fibroblasts . gelatin zymography and western immunoblot analysis for mmp - 9 were performed on the coculture cell media and the single cell media . inhibition of fibroblast - mediated breast tumor cell growth by an anti - tsp - 1 or an anti - mmp - 9 antibody was evaluated using a modified boyden chamber . results coculture experiments showed an increased production of mmp - 9 when compared with breast cancer single cell culture or fibroblast single cell culture experiments as demonstrated by zymography and western immunoblot analysis . fibroblast - stimulated mmp - 9 production was similar with il - 1 - stimulated mmp - 9 production . anti - st - 1 antibody and anti - mmp - 9 antibody inhibited fibroblast - stimulated tumor cell growth to 30 % and 26 % of controls , respectively (p $<$. 05) . conclusions fibroblasts may regulate breast cancer cell proliferation by promoting tumor mmp - 9 production $F(s_{\text{adv}}, l =$"<multilabel>"$) = 0.99$ **Cos.** = -0.2 $MedSTS = 0.72$	background caveolin - 1 (sdf - 1) promotes breast cancer cell invasion of collagen by upregulating matrix metalloproteinase - 9 (mmp - 9) production . autocrine igf - 1 may play a role in regulating tumor cell invasion . we hypothesize that caf promote breast cancer cell invasion by upregulating the production of mmp - 9 through tsp - 1 . methods mda - mb - 231 human breast carcinoma cells were grown alone or in coculture with human fibroblasts . gelatin zymography and western immunoblot analysis for mmp - 9 were performed on the coculture cell media and the single cell media . inhibition of coculture - mediated breast tumor cell invasion by an anti - icam - 1 or an anti - mmp - 9 antibody was evaluated using a modified boyden chamber . results coculture experiments showed an increased production of mmp - 9 when compared with breast cancer single cell culture or fibroblast single cell culture experiments as demonstrated by zymography and western immunoblot analysis . fibroblast - stimulated mmp - 9 production was consistent with sdf - 1 - upregulated mmp - 9 production . anti - timp - 1 antibody and anti - mmp - 9 antibody inhibited fibroblast - stimulated tumor cell invasion to 30 % and 26 % of controls , respectively (p $<$. 05) . conclusions fibroblast may regulate breast cancer cell invasion by promoting tumor mmp - 9 production $F(s_{\text{adv}}, l =$"<multilabel>"$) = 1.0$ **Cos.** = 0.03 $MedSTS = 0.67$	background endothelin - 1 (mmp - 1) promotes breast cancer cell invasion of collagen by upregulating matrix metalloproteinase - 9 (mmp - 9) production . hypothesis spp - 1 may play a role in regulating tumor cell invasion . we hypothesize that adipocytes promote breast cancer cell migration by upregulating the production of mmp - 9 through tsp - 1 . methods mda - mb - 231 human breast carcinoma cells were grown alone or in coculture with human fibroblasts . gelatin zymography and western immunoblot analysis for mmp - 9 were performed on the coculture cell media and the single cell media . inhibition of adipocyte - mediated breast tumor cell invasion by an anti - pai - 1 or an anti - mmp - 9 antibody was evaluated using a modified boyden chamber . results coculture experiments showed an increased production of mmp - 9 when compared with breast cancer single cell culture or fibroblast single cell culture experiments as demonstrated by zymography and western immunoblot analysis . fibroblast - stimulated mmp - 9 production was similar with il - 1 - stimulated mmp - 9 production . anti - mcp - 1 antibody and anti - mmp - 9 antibody inhibited fibroblast - stimulated tumor cell invasion to 30 % and 26 % of controls , respectively (p $<$. 05) . conclusions fibroblast may regulate breast cancer cell invasion by promoting tumor mmp - 9 production $F(s_{\text{adv}}, l =$"<multilabel>"$) = 1.0$ **Cos.** = 0.21 $MedSTS = 0.66$
Original	although regulatory t cells (t (regs)) are known to suppress self - reactive autoimmune responses , their role during t cell responses to nonself antigens is not well understood . we show that t (regs) play a critical role during the priming of immune responses in mice . t (reg) depletion induced the activation and expansion of a population of low - avidity cd8 (+) t cells because of overproduction of ccl - 3 / 4 / 5 chemokines , which stabilized the interactions between antigen - presenting dendritic cells and low - avidity t cells . in the absence of t (regs) , the avidity of the primary immune response was impaired , which resulted in reduced memory to listeria monocytogenes . these results suggest that t (regs) are important regulators of the homeostasis of cd8 (+) t cell priming and play a critical role in the induction of high - avidity primary responses and effective memory . $F(s, l =$"<multilabel>"$) = 1.0$	although regulatory t cells (t (regs)) are known to suppress self - reactive autoimmune responses , their role during t cell responses to nonself antigens is not well understood . we show that t (regs) play a critical role during the priming of immune responses in mice . t (reg) depletion induced the activation and expansion of a population of low - avidity cd8 (+) t cells because of overproduction of ccl - 3 / 4 / 5 chemokines , which stabilized the interactions between antigen - presenting dendritic cells and low - avidity t cells . in the absence of t (regs) , the avidity of the primary immune response was impaired , which resulted in reduced memory to listeria monocytogenes . these results suggest that t (regs) are important regulators of the homeostasis of cd8 (+) t cell priming and play a critical role in the induction of high - avidity primary responses and effective memory . $F(s, l =$"<multilabel>"$) = 1.0$	although regulatory t cells (t (regs)) are known to suppress self - reactive autoimmune responses , their role during t cell responses to nonself antigens is not well understood . we show that t (regs) play a critical role during the priming of immune responses in mice . t (reg) depletion induced the activation and expansion of a population of low - avidity cd8 (+) t cells because of overproduction of ccl - 3 / 4 / 5 chemokines , which stabilized the interactions between antigen - presenting dendritic cells and low - avidity t cells . in the absence of t (regs) , the avidity of the primary immune response was impaired , which resulted in reduced memory to listeria monocytogenes . these results suggest that t (regs) are important regulators of the homeostasis of cd8 (+) t cell priming and play a critical role in the induction of high - avidity primary responses and effective memory . $F(s, l =$"<multilabel>"$) = 1.0$

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	although natural effector tregs (t (regs)) are known to suppress self - reactive immune responses , their role during t cell responses to nonself antigens is not well understood . we show that t (regs) play a critical role during the priming of primary responses in mice . t (reg) depletion induced the activation and expansion of a population of low - avidity cd8 (+) t cells because of overproduction of ccl - 3 / 4 / 5 cytokines , which stabilized the interactions between antigen - presenting target cells and low - avidity t cells . in the absence of t (regs) , the avidity of the primary antibody response was impaired , which resulted in reduced memory to listeria monocytogenes . these results suggest that t (regs) are important regulators of the homeostasis of cd8 (+) t cell priming and play a critical role in the induction of high - avidity primary responses and effective memory . $F(\boldsymbol{s}_{\text{adv}}, l =$ "<multilabel>"$) = 1.0$ **Cos.** = -0.09 $MedSTS$ = 0.95	although reg t cells (t (regs)) are known to suppress self - reactive inflammatory responses , their role during t cell responses to nonself antigens is not well understood . we show that t (regs) play a critical role during the priming of immune responses in mice . t (reg) depletion induced the activation and expansion of a population of low - activated cd8 (+) t cells because of overproduction of ccl - 3 / 4 / 5 expression , which stabilized the interactions between antigen - presenting dendritic cells and low - threshold t cells . in the absence of t (regs) , the priming of the primary immune response was impaired , which resulted in reduced memory to listeria monocytogenes . these results suggest that t (regs) are important regulators of the outcome of cd8 (+) t cell priming and play a critical role in the induction of high - affinity primary responses and effective memory . $F(\boldsymbol{s}_{\text{adv}}, l =$ "<multilabel>"$) = 1.0$ **Cos.** = 0.12 $MedSTS$ = 0.96	although reg t cells (t (regs)) are known to suppress self - reactive immune responses , their role during t cell responses to nonself antigens is not well understood . we show that t (regs) play a critical role during the priming of immune responses in mice . t (reg) depletion induced the activation and expansion of a population of low - avidity cd8 (+) t cells because of overproduction of gata - 3 / 4 / 5 antibodies , which stabilized the interactions between antigen - presenting accessory cells and low - avidity t cells . in the absence of t (regs) , the priming of the primary immune response was impaired , which resulted in reduced memory to listeria monocytogenes . these results suggest that t (regs) are important regulators of the efficiency of cd8 (+) t cell priming and play a critical role in the induction of high - affinity primary responses and effective memory . $F(\boldsymbol{s}_{\text{adv}}, l =$ "<multilabel>"$) = 1.0$ **Cos.** = 0.49 $MedSTS$ = 0.95
Original	class switch recombination (csr) in b lymphocytes is initiated by introduction of multiple dna double - strand breaks (dsbs) into switch (s) regions that flank immunoglobulin heavy chain (igh) constant region exons . csr is completed by joining a dsb in the donor s mu to a dsb in a downstream acceptor s region (e . g . , s gamma1) by end - joining . in normal cells , many csr junctions are mediated by classical nonhomologous end - joining (c - nhej) , which employs the ku70 / 80 complex for dsb recognition and xrcc4 / dna ligase 4 for ligation . alternative end - joining (a - ej) mediates csr , at reduced levels , in the absence of c - nhej , even in combined absence of ku70 and ligase 4 , demonstrating an a - ej pathway totally distinct from c - nhej . multiple dsbs are introduced into s mu during csr , with some being rejoined or joined to each other to generate internal switch deletions (isds) . in addition , s - region dsbs can be joined to other chromosomes to generate translocations , the level of which is increased by absence of a single c - nhej component (e . g . , xrcc $F(\boldsymbol{s}, l =$ "<multilabel>"$) = 1.0$	class switch recombination (csr) in b lymphocytes is initiated by introduction of multiple dna double - strand breaks (dsbs) into switch (s) regions that flank immunoglobulin heavy chain (igh) constant region exons . csr is completed by joining a dsb in the donor s mu to a dsb in a downstream acceptor s region (e . g . , s gamma1) by end - joining . in normal cells , many csr junctions are mediated by classical nonhomologous end - joining (c - nhej) , which employs the ku70 / 80 complex for dsb recognition and xrcc4 / dna ligase 4 for ligation . alternative end - joining (a - ej) mediates csr , at reduced levels , in the absence of c - nhej , even in combined absence of ku70 and ligase 4 , demonstrating an a - ej pathway totally distinct from c - nhej . multiple dsbs are introduced into s mu during csr , with some being rejoined or joined to each other to generate internal switch deletions (isds) . in addition , s - region dsbs can be joined to other chromosomes to generate translocations , the level of which is increased by absence of a single c - nhej component (e . g . , xrcc $F(\boldsymbol{s}, l =$ "<multilabel>"$) = 1.0$	class switch recombination (csr) in b lymphocytes is initiated by introduction of multiple dna double - strand breaks (dsbs) into switch (s) regions that flank immunoglobulin heavy chain (igh) constant region exons . csr is completed by joining a dsb in the donor s mu to a dsb in a downstream acceptor s region (e . g . , s gamma1) by end - joining . in normal cells , many csr junctions are mediated by classical nonhomologous end - joining (c - nhej) , which employs the ku70 / 80 complex for dsb recognition and xrcc4 / dna ligase 4 for ligation . alternative end - joining (a - ej) mediates csr , at reduced levels , in the absence of c - nhej , even in combined absence of ku70 and ligase 4 , demonstrating an a - ej pathway totally distinct from c - nhej . multiple dsbs are introduced into s mu during csr , with some being rejoined or joined to each other to generate internal switch deletions (isds) . in addition , s - region dsbs can be joined to other chromosomes to generate translocations , the level of which is increased by absence of a single c - nhej component (e . g . , xrcc $F(\boldsymbol{s}, l =$ "<multilabel>"$) = 1.0$

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	class switch switching (csr) in b lymphocytes is initiated by introduction of multiple intras - - dna ends (dsbs) into switch (s) regions that flank immunoglobulin heavy chain (hc) constant region segments . csr is completed by joining a dsb in the donor s mu to a dsb in a downstream acceptor s region (e . g . , s gamma1) by end - joining . in normal cells , many csr junctions are mediated by classical nonhomologous end - joining (c - nhej) , which employs the ku70 / 80 complex for dsb recognition and xrcc4 / dna ligase 4 for ligation . alternative end - joining (a - ej) mediates csr , at reduced levels , in the absence of c - nhej , even in combined absence of ku70 and ligase 4 , demonstrating an a - ej pathway totally distinct from c - nhej . multiple dsbs are introduced into s mu during csr , with some being rejoined or joined to each other to generate internal satellite dnas (isds) . in addition , s - region dsbs can be joined to other ends to generate is , the level of which is increased by absence of a single c - nhej component (e . g . , xrcc $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 1.0 **Cos.** = 0.37 $MedSTS$ = 0.57	class switch recombination (csr) in b lymphocytes is initiated by introduction of multiple dsdna double - strand ends (dsbs) into switch (s) regions that flank immunoglobulin heavy chain (h) constant region exons . csr is completed by joining a dsb in the donor c mu to a dsb in a downstream acceptor s region (e . g . , s gamma1) by end - joining . in normal cells , many csr junctions are mediated by classical nonhomologous end - joining (c - nhej) , which employs the ku70 / 80 complex for dsb recognition and xrcc4 / dna ligase 4 for ligation . alternative end - joining (a - ej) mediates csr , at reduced levels , in the absence of c - nhej , even in combined absence of ku70 and ligase 4 , demonstrating an a - ej pathway totally distinct from c - nhej . multiple dsbs are introduced into switch dna during csr , with some being rejoined or joined to each other to generate insertion switch deletions (isds) . in addition , switch - region dsbs can be joined to other together to generate indels , the level of which is increased by absence of a single c - nhej component (e . g . , , $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 1.0 **Cos.** = 0.5 $MedSTS$ = 0.54	class switching diversification (csr) in b lineage is initiated by introduction of multiple , double - strand breaks (dsbs) into switch (s) regions that flank immunoglobulin heavy chain (h) constant region exons . csr is completed by joining a dsb in the donor s mu to a dsb in a downstream acceptor s region (e . g . , s gamma1) by end - joining . in normal cells , many csr junctions are mediated by classical nonhomologous end - joining (c - nhej) , which employs the ku70 / 80 complex for dsb recognition and xrcc4 / dna ligase 4 for ligation . alternative end - joining (a - ej) mediates csr , at reduced levels , in the absence of c - nhej , even in combined absence of ku70 and ligase 4 , demonstrating an a - ej pathway totally distinct from c - nhej . multiple dsbs are introduced into the mu during csr , with some being rejoined or joined to each other to generate internal recombination domains (isds) . in addition , switch - region dsbs can be joined to other dsbs to generate chimeras , the level of which is increased by absence of a single c - nhej component (e . g . , xr $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 1.0 **Cos.** = 0.7 $MedSTS$ = 0.55
Original	epidemiological studies suggest that dietary polyunsaturated fatty acids (pufa) may influence breast cancer progression and prognosis . in order to study potential mechanisms of action of fatty acid modulation of tumor growth , we studied , in vitro , the influence of n - 3 and n - 6 fatty acids on proliferation , cell cycle , differentiation and apoptosis of mcf - 7 human breast cancer cells . both eicosapentaenoic acid (epa) and docosahexaenoic acid (dha) inhibited the mcf - 7 cell growth by 30 % and 54 % , respectively , while linoleic acid (la) had no effect and arachidonic acid (aa) inhibited the cell growth by 30 % (p < 0 . 05) . the addition of vitamin e (10um) to cancer cells slightly restored cell growth . the incubation of mcf - 7 cells with pufas did not alter the cell cycle parameters or induce cell apoptosis . however , the growth inhibitory effects of epa , dha and aa were associated with cell differentiation as indicated by positive oil - red - o staining of the cells . lipid droplet accumulation was increased by 65 % , 30 % and 15 % in the presence of dha , epa and aa , respectively ; (p < 0 . 05) . these observations suggest that fatty acids may influence cellular processes at a molecular level , capable of modulating breast cancer cell growth $F(\boldsymbol{s}, l =$"<multilabel>") = 1.0	epidemiological studies suggest that dietary polyunsaturated fatty acids (pufa) may influence breast cancer progression and prognosis . in order to study potential mechanisms of action of fatty acid modulation of tumor growth , we studied , in vitro , the influence of n - 3 and n - 6 fatty acids on proliferation , cell cycle , differentiation and apoptosis of mcf - 7 human breast cancer cells . both eicosapentaenoic acid (epa) and docosahexaenoic acid (dha) inhibited the mcf - 7 cell growth by 30 % and 54 % , respectively , while linoleic acid (la) had no effect and arachidonic acid (aa) inhibited the cell growth by 30 % (p < 0 . 05) . the addition of vitamin e (10um) to cancer cells slightly restored cell growth . the incubation of mcf - 7 cells with pufas did not alter the cell cycle parameters or induce cell apoptosis . however , the growth inhibitory effects of epa , dha and aa were associated with cell differentiation as indicated by positive oil - red - o staining of the cells . lipid droplet accumulation was increased by 65 % , 30 % and 15 % in the presence of dha , epa and aa , respectively ; (p < 0 . 05) . these observations suggest that fatty acids may influence cellular processes at a molecular level , capable of modulating breast cancer cell growth $F(\boldsymbol{s}, l =$"<multilabel>") = 1.0	epidemiological studies suggest that dietary polyunsaturated fatty acids (pufa) may influence breast cancer progression and prognosis . in order to study potential mechanisms of action of fatty acid modulation of tumor growth , we studied , in vitro , the influence of n - 3 and n - 6 fatty acids on proliferation , cell cycle , differentiation and apoptosis of mcf - 7 human breast cancer cells . both eicosapentaenoic acid (epa) and docosahexaenoic acid (dha) inhibited the mcf - 7 cell growth by 30 % and 54 % , respectively , while linoleic acid (la) had no effect and arachidonic acid (aa) inhibited the cell growth by 30 % (p < 0 . 05) . the addition of vitamin e (10um) to cancer cells slightly restored cell growth . the incubation of mcf - 7 cells with pufas did not alter the cell cycle parameters or induce cell apoptosis . however , the growth inhibitory effects of epa , dha and aa were associated with cell differentiation as indicated by positive oil - red - o staining of the cells . lipid droplet accumulation was increased by 65 % , 30 % and 15 % in the presence of dha , epa and aa , respectively ; (p < 0 . 05) . these observations suggest that fatty acids may influence cellular processes at a molecular level , capable of modulating breast cancer cell growth $F(\boldsymbol{s}, l =$"<multilabel>") = 0.94

	Vanilla	Adversarial	FAR-IG
Adversarial	epidemiological studies suggest that dietary polyunsaturated fatty acids (pufa) may influence breast cancer progression and prognosis . in order to study potential mechanisms of action of fatty acid modulation of tumor growth , we studied , in vitro , the influence of n - 3 and n - 6 fatty acids on proliferation , cell death , migration and apoptosis of mcf - 7 human breast cancer cells . both eicosapentaenoic acid (epa) and docosahexaenoic acid (dha) inhibited the mcf - 7 cell growth by 30 % and 54 % , respectively , while linoleic acid (la) had no effect and arachidonic acid (aa) inhibited the cell growth by 30 % (p < 0 . 05) . the addition of exogenous e (10um) to cancer cells slightly restored cell growth . the incubation of mcf - 7 cells with pufas did not alter the - biochemical features or their cell transformation . however , the growth inhibitory effects of epa , dha and aa were associated with cell proliferation as indicated by positive oil - red - o staining of the cells . tri cholesterol uptake was increased by 65 % , 30 % and 15 % in the presence of dha , epa and aa , respectively ; (p < 0 . 05) . these observations suggest that fatty acids may influence cellular processes at a molecular level , capable of modulating breast cancer cell growth $F(s_{\text{adv}}, l =$ "<multilabel>") = 0.99 **Cos.** = -0.08 $MedSTS$ = 0.9	epidemiological studies suggest that dietary polyunsaturated fatty acids (pufa) may influence breast cancer progression and prognosis . in order to study potential mechanisms of action of fatty acid modulation of tumor growth , we studied , in vitro , the influence of n - 3 and n - 6 fatty acids on proliferation , cell cycle , migration and invasion of mcf - 7 human breast cancer cells . both eicosapentaenoic acid (epa) and docosahexaenoic acid (dha) inhibited the mcf - 7 cell growth by 30 % and 54 % , respectively , while linoleic acid (la) had no effect and linoleic acid (aa) inhibited the cell growth by 30 % (p < 0 . 05) . the addition of prostaglandin e (10um) to cancer cells slightly restored cell growth . the incubation of mcf - 7 cells with pufas did not alter the cell proliferation parameters or induce cell transformation . however , the growth inhibitory effects of epa , dha and aa were related with cell proliferation as indicated by positive nuclear - red - o labeling of the cells . lipid apoptotic formation was increased by 65 % , 30 % and 15 % in the presence of dha , epa and aa , respectively ; (p < 0 . 05) . these observations suggest that fatty acids may influence cellular processes at a molecular level , capable of modulating breast cancer cell growth $F(s_{\text{adv}}, l =$ "<multilabel>") = 0.97 **Cos.** = 0.08 $MedSTS$ = 0.81	epidemiological studies suggest that dietary polyunsaturated fatty acids (pufa) may influence breast cancer progression and prognosis . in order to study potential mechanisms of action of fatty acid modulation of tumor growth , we studied , in vitro , the influence of n - 3 and n - 6 fatty acids on proliferation , cell cycle , proliferation and morphology of mcf - 7 human breast cancer cells . both epa acid (epa) and docosahexaenoic acid (dha) inhibited the mcf - 7 cell growth by 30 % and 54 % , respectively , while linoleic acid (la) had no effect and acetic acid (aa) inhibited the cell growth by 30 % (p < 0 . 05) . the addition of prostaglandin e (10um) to cancer cells slightly restored cell growth . the incubation of mcf - 7 cells with pufa did not alter the cell proliferation parameters or induce cell differentiation . however , the growth inhibitory effects of epa , dha and aa were associated with cell proliferation as indicated by positive methyl - red - o staining of the cells . lipid staining formation was increased by 65 % , 30 % and 15 % in the presence of dha , epa and aa , respectively ; (p < 0 . 05) . these observations suggest that fatty acids may influence cellular processes at a molecular level , capable of modulating breast cancer cell growth $F(s_{\text{adv}}, l =$ "<multilabel>") = 0.78 **Cos.** = 0.2 $MedSTS$ = 0.78
Original	the warburg effect describes a heightened propensity of tumor cells to produce lactic acid in the presence or absence of o (2) . a generally held notion is that the warburg effect is related to energy . using whole - genome , proteomic maldi - tof - ms and metabolite analysis , we investigated the warburg effect in malignant neuroblastoma n2a cells . the findings show that the warburg effect serves a functional role in regulating acidic pericellular ph (phe) , which is mediated by metabolic inversion or a fluctuating dominance between glycolytic - rate substrate level phosphorylation (slp) and mitochondrial (mt) oxidative phosphorylation (oxphos) to control lactic acid production . the results also show that an alkaline phe caused an elevation in slp / oxphos ratio (approximately 98 % slp / oxphos) ; while the ratio was approximately 56 % at neutral phe and approximately 93 % in acidic phe . acidic phe paralleled greater expression of mitochondrial biogenesis and oxphos genes , such as complex iii - v (uqcr10 , atp5 and cox7c) , mt fmc1 , romo1 , tmem 173 , tomm6 , aldehyde dehydrogenase , mt sod2 mt biogenesis component ppar - \ u03b3 co - activator 1 adjunct to loss of mt fission (mff) . $F(s, l =$ "<multilabel>") = 1.0	the warburg effect describes a heightened propensity of tumor cells to produce lactic acid in the presence or absence of o (2) . a generally held notion is that the warburg effect is related to energy . using whole - genome , proteomic maldi - tof - ms and metabolite analysis , we investigated the warburg effect in malignant neuroblastoma n2a cells . the findings show that the warburg effect serves a functional role in regulating acidic pericellular ph (phe) , which is mediated by metabolic inversion or a fluctuating dominance between glycolytic - rate substrate level phosphorylation (slp) and mitochondrial (mt) oxidative phosphorylation (oxphos) to control lactic acid production . the results also show that an alkaline phe caused an elevation in slp / oxphos ratio (approximately 98 % slp / oxphos) ; while the ratio was approximately 56 % at neutral phe and approximately 93 % in acidic phe . acidic phe paralleled greater expression of mitochondrial biogenesis and oxphos genes , such as complex iii - v (uqcr10 , atp5 and cox7c) , mt fmc1 , romo1 , tmem 173 , tomm6 , aldehyde dehydrogenase , mt sod2 mt biogenesis component ppar - \ u03b3 co - activator 1 adjunct to loss of mt fission (mff) . $F(s, l =$ "<multilabel>") = 1.0	the warburg effect describes a heightened propensity of tumor cells to produce lactic acid in the presence or absence of o (2) . a generally held notion is that the warburg effect is related to energy . using whole - genome , proteomic maldi - tof - ms and metabolite analysis , we investigated the warburg effect in malignant neuroblastoma n2a cells . the findings show that the warburg effect serves a functional role in regulating acidic pericellular ph (phe) , which is mediated by metabolic inversion or a fluctuating dominance between glycolytic - rate substrate level phosphorylation (slp) and mitochondrial (mt) oxidative phosphorylation (oxphos) to control lactic acid production . the results also show that an alkaline phe caused an elevation in slp / oxphos ratio (approximately 98 % slp / oxphos) ; while the ratio was approximately 56 % at neutral phe and approximately 93 % in acidic phe . acidic phe paralleled greater expression of mitochondrial biogenesis and oxphos genes , such as complex iii - v (uqcr10 , atp5 and cox7c) , mt fmc1 , romo1 , tmem 173 , tomm6 , aldehyde dehydrogenase , mt sod2 mt biogenesis component ppar - \ u03b3 co - activator 1 adjunct to loss of mt fission (mff) . $F(s, l =$ "<multilabel>") = 1.0

	Vanilla	Adversarial	FAR-IG
Adversarial	the pasteur effect describes a heightened propensity of living cells to produce organic acid in the presence or absence of o (2) . a generally held notion is that the warburg effect is related to energy . using whole - genome , proteomic maldi - tof - ms and metabolite analysis , we investigated the ph effect in malignant neuronal n2a cells . the findings show that the hall effect serves a functional role in regulating acidic ph acidosis (phe) , which is mediated by metabolic inversion or a fluctuating dominance between steady - rate substrate level phosphorylation (slp) and mitochondrial (mt) oxidative phosphorylation (oxphos) to control extracellular acid production . the results also show that an alkaline phe caused an elevation in slp / oxphos ratio (approximately 98 % slp / oxphos) ; while the ratio was approximately 56 % at neutral phe and approximately 93 % in acidic phe . acidic phe paralleled greater expression of mitochondrial biogenesis and oxphos genes , such as complex iii - v (uqcr10 , atp5 and cox7c) , mt fmc1 , romo1 , tmem 173 , tomm6 , aldehyde dehydrogenase , mt sod2 mt biogenesis component ppar - \ u03b3 co - activator 1 adjunct to loss of mt fluorescence (mff) . $F(\boldsymbol{s}_{\mathrm{adv}},\ l=$“<multilabel>”$) = 1.0$ **Cos.** = 0.2 $MedSTS$ = 0.59	the warburg effect describes a heightened propensity of tumor cells to produce lactate acid in the presence or absence of o (2) . a generally held notion is that the warburg effect is related to energy . using whole - genome , proteomic maldi - tof - ms and metabolite analysis , we investigated the warburg effect in malignant glioma n2a cells . the findings show that the warburg effect serves a functional role in regulating tumor energy production (ph) , which is mediated by metabolic inversion or a fluctuating dominance between low - rate substrate level phosphorylation (slp) and mitochondrial (mt) oxidative phosphorylation (oxphos) to control lactic acid production . the results also show that an alkaline phe caused an elevation in slp / oxphos ratio (approximately 98 % slp / oxphos) ; while the ratio was approximately 56 % at neutral phe and approximately 93 % in alkaline phe . acidic phe showed greater expression of organelle biogenesis and oxphos genes , such as complex iii - v (uqcr10 , atp5 and cox7c) , mt fmc1 , romo1 , tmem 173 , tomm6 , aldehyde dehydrogenase , mt sod2 mt biogenesis component ppar - \ u03b3 co - activator 1 adjunct to loss of mitochondrial fission (mff) . $F(\boldsymbol{s}_{\mathrm{adv}},\ l=$“<multilabel>”$) = 1.0$ **Cos.** = 0.43 $MedSTS$ = 0.61	the warburg effect describes a heightened propensity of malignant cells to produce lactic acid in the presence or absence of o (2) . a generally held notion is that the warburg effect is related to energy . using whole - genome , proteomic maldi - tof - ms and metabolite analysis , we investigated the warburg effect in malignant mouse n2a cells . the findings show that the warburg effect serves a functional role in regulating tumor pericellular acidification (phe) , which is mediated by metabolic inversion or a fluctuating dominance between steady - rate substrate level phosphorylation (slp) and matrix (mt) oxidative phosphorylation (oxphos) to control lactic acid production . the results also show that an alkaline phe caused an elevation in slp / oxphos ratio (approximately 98 % slp / oxphos) ; while the ratio was approximately 56 % at neutral phe and approximately 93 % in acidic phe . acidic phe paralleled greater expression of organelle biogenesis and oxphos genes , such as complex iii - v (uqcr10 , atp5 and cox7c) , tom fmc1 , romo1 , tom 173 , tomm6 , aldehyde dehydrogenase , mt sod2 mt biogenesis component , - \ u03b3 co - activator 1 adjunct to loss of mitochondria fission (mff) . $F(\boldsymbol{s}_{\mathrm{adv}},\ l=$“<multilabel>”$) = 1.0$ **Cos.** = 0.56 $MedSTS$ = 0.64
Original	to investigate whether altered energy metabolism induces the warburg effect and results in tumor malignancy , the respiratory enzyme citrate synthase (cs) was examined , silenced , and the effects analyzed . in human cervical carcinoma cells , rnai - mediated cs knockdown induced morphological changes characteristic of the epithelial - mesenchymal transition (emt) . this switch accelerated cancer cell metastasis and proliferation in in vitro assays and in vivo tumor xenograft models . notably , cs knockdown cells exhibited severe defects in respiratory activity and marked decreases in atp production , but great increases in glycolytic metabolism . this malignant progression was due to activation of emt - related regulators ; altered energy metabolism resulted from deregulation of the p53 / tigar and sco2 pathways . this phenotypic change was completely reversed by p53 reactivation via treatment with proteasome inhibitor mg132 or co - knockdown of e3 ligase hdm2 and partially suppressed by atp treatment . this study directly links the warburg effect to tumor malignancy via induction of the emt phenotype . $F(\boldsymbol{s},\ l=$“<multilabel>”$) = 1.0$	to investigate whether altered energy metabolism induces the warburg effect and results in tumor malignancy , the respiratory enzyme citrate synthase (cs) was examined , silenced , and the effects analyzed . in human cervical carcinoma cells , rnai - mediated cs knockdown induced morphological changes characteristic of the epithelial - mesenchymal transition (emt) . this switch accelerated cancer cell metastasis and proliferation in in vitro assays and in vivo tumor xenograft models . notably , cs knockdown cells exhibited severe defects in respiratory activity and marked decreases in atp production , but great increases in glycolytic metabolism . this malignant progression was due to activation of emt - related regulators ; altered energy metabolism resulted from deregulation of the p53 / tigar and sco2 pathways . this phenotypic change was completely reversed by p53 reactivation via treatment with proteasome inhibitor mg132 or co - knockdown of e3 ligase hdm2 and partially suppressed by atp treatment . this study directly links the warburg effect to tumor malignancy via induction of the emt phenotype . $F(\boldsymbol{s},\ l=$“<multilabel>”$) = 1.0$	to investigate whether altered energy metabolism induces the warburg effect and results in tumor malignancy , the respiratory enzyme citrate synthase (cs) was examined , silenced , and the effects analyzed . in human cervical carcinoma cells , rnai - mediated cs knockdown induced morphological changes characteristic of the epithelial - mesenchymal transition (emt) . this switch accelerated cancer cell metastasis and proliferation in in vitro assays and in vivo tumor xenograft models . notably , cs knockdown cells exhibited severe defects in respiratory activity and marked decreases in atp production , but great increases in glycolytic metabolism . this malignant progression was due to activation of emt - related regulators ; altered energy metabolism resulted from deregulation of the p53 / tigar and sco2 pathways . this phenotypic change was completely reversed by p53 reactivation via treatment with proteasome inhibitor mg132 or co - knockdown of e3 ligase hdm2 and partially suppressed by atp treatment . this study directly links the warburg effect to tumor malignancy via induction of the emt phenotype . $F(\boldsymbol{s},\ l=$“<multilabel>”$) = 1.0$

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	to investigate whether altered energy metabolism induces the bystander effect and results in tumor malignancy , the respiratory enzyme citrate synthase (cs) was examined , silenced , and the effects analyzed . in human cervical carcinoma cells , rnai - mediated cs knockdown induced morphological changes characteristic of the epithelial - mesenchymal transition (eoc) . this switch accelerated cancer cell viability and proliferation in in vitro assays and in vivo tumor xenograft models . notably , cs knockdown cells exhibited severe defects in respiratory activity and marked decreases in atp production , but great increases in energy metabolism . this malignant progression was due to activation of stemness - related regulators ; altered energy metabolism resulted from deregulation of the mtor / tigar and sco2 pathways . this phenotypic change was completely reversed by p53 reactivation via treatment with proteasome inhibitor mg132 or co - knockdown of e3 ligase hdm2 and partially suppressed by atp treatment . this study directly links the reprogramming effect to tumor malignancy via induction of the transition phenotype . $F(\boldsymbol{s}_{\text{adv}}, l =$ "<multilabel>"$) = 0.94$ **Cos.** = 0.13 $MedSTS = 0.8$	to investigate whether altered energy metabolism induces the warburg effect and results in tumor malignancy , the respiratory enzyme citrate synthase (cs) was examined , silenced , and the effects analyzed . in human cervical carcinoma cells , rnai - mediated cs knockdown induced morphological changes characteristic of the epithelial - carcinoma transition (eoc) . this switch accelerated cancer cell metastasis and proliferation in in vitro assays and in vivo tumor xenograft models . notably , cs knockdown cells exhibited severe defects in respiratory activity and marked decreases in atp production , but great increases in lipid metabolism . this malignant progression was due to activation of stemness - related regulators ; altered energy metabolism resulted from deregulation of the mtor / ampk and sco2 pathways . this phenotypic change was completely reversed by p53 reactivation via treatment with mg132 inhibitor mg132 or co - knockdown of e3 ligase hdm2 and partially suppressed by atp treatment . this study directly links the warburg effect to tumor malignancy via induction of the transition phenotype . $F(\boldsymbol{s}_{\text{adv}}, l =$ "<multilabel>"$) = 0.8$ **Cos.** = 0.5 $MedSTS = 0.77$	to investigate whether altered energy metabolism induces the warburg effect and results in tumor metastasis , the respiratory enzyme citrate synthase (cs) was examined , silenced , and the effects analyzed . in human cervical carcinoma cells , rnai - mediated cs knockdown induced morphological changes characteristic of the epithelial - mesenchyme transformation (emt) . this switch accelerated cancer cell metastasis and proliferation in in vitro assays and in vivo tumor xenograft models . notably , cs knockdown cells exhibited severe defects in respiratory activity and marked decreases in atp production , but great increases in energy metabolism . this malignant progression was due to activation of energy - related regulators ; altered energy metabolism resulted from deregulation of the p53 / tigar and sco2 pathways . this phenotypic change was completely reversed by p53 reactivation via treatment with proteasomal inhibitor mg132 or co - knockdown of e3 ligase hdm2 and partially suppressed by atp treatment . this study directly links the warburg effect to tumor metastasis via induction of the mesenchymal phenotype . $F(\boldsymbol{s}_{\text{adv}}, l =$ "<multilabel>"$) = 1.0$ **Cos.** = 0.63 $MedSTS = 0.82$
Original	missense substitutions of uncertain clinical significance in the brca1 gene are a vexing problem in genetic counseling for women who have a family history of breast cancer . in this study , we evaluated the functions of 29 missense substitutions of brca1 in two dna repair pathways . repair of double - strand breaks by homology - directed recombination (hdr) had been previously analyzed for 16 of these brca1 variants , and 13 more variants were analyzed in this study . all 29 variants were also analyzed for function in double - strand break repair by the single - strand annealing (ssa) pathway . we found that among the pathogenic mutations in brca1 , all were defective for dna repair by either pathway . the hdr assay was accurate because all pathogenic mutants were defective for hdr , and all nonpathogenic variants were fully functional for hdr . repair by ssa accurately identified pathogenic mutants , but several nonpathogenic variants were scored as defective or partially defective . these results indicated that specific amino acid residues of the brca1 protein have different effects in the two related dna repair pathways , and these results validate the hdr assay as highly correlative with brca1 - associated breast cancer . $F(\boldsymbol{s}, l =$ "<multilabel>"$) = 1.0$	missense substitutions of uncertain clinical significance in the brca1 gene are a vexing problem in genetic counseling for women who have a family history of breast cancer . in this study , we evaluated the functions of 29 missense substitutions of brca1 in two dna repair pathways . repair of double - strand breaks by homology - directed recombination (hdr) had been previously analyzed for 16 of these brca1 variants , and 13 more variants were analyzed in this study . all 29 variants were also analyzed for function in double - strand break repair by the single - strand annealing (ssa) pathway . we found that among the pathogenic mutations in brca1 , all were defective for dna repair by either pathway . the hdr assay was accurate because all pathogenic mutants were defective for hdr , and all nonpathogenic variants were fully functional for hdr . repair by ssa accurately identified pathogenic mutants , but several nonpathogenic variants were scored as defective or partially defective . these results indicated that specific amino acid residues of the brca1 protein have different effects in the two related dna repair pathways , and these results validate the hdr assay as highly correlative with brca1 - associated breast cancer . $F(\boldsymbol{s}, l =$ "<multilabel>"$) = 1.0$	missense substitutions of uncertain clinical significance in the brca1 gene are a vexing problem in genetic counseling for women who have a family history of breast cancer . in this study , we evaluated the functions of 29 missense substitutions of brca1 in two dna repair pathways . repair of double - strand breaks by homology - directed recombination (hdr) had been previously analyzed for 16 of these brca1 variants , and 13 more variants were analyzed in this study . all 29 variants were also analyzed for function in double - strand break repair by the single - strand annealing (ssa) pathway . we found that among the pathogenic mutations in brca1 , all were defective for dna repair by either pathway . the hdr assay was accurate because all pathogenic mutants were defective for hdr , and all nonpathogenic variants were fully functional for hdr . repair by ssa accurately identified pathogenic mutants , but several nonpathogenic variants were scored as defective or partially defective . these results indicated that specific amino acid residues of the brca1 protein have different effects in the two related dna repair pathways , and these results validate the hdr assay as highly correlative with brca1 - associated breast cancer . $F(\boldsymbol{s}, l =$ "<multilabel>"$) = 1.0$

	Vanilla	Adversarial	FAR-IG
Adversarial	missense substitutions of unknown clinical significance in the brca1 gene are a vexing problem in risk counseling for women who have a family history of breast cancer . in this study , we evaluated the functions of 29 missense variants of brca1 in two dna repair pathways . repair of two - strand breaks by homology - directed recombination (hdr) had been previously analyzed for 16 of these brca1 variants , and 13 more variants were analyzed in this study . all 29 variants were also analyzed for function in single - strand break repair by the single - strand annealing (ssa) pathway . we found that among the 29 variants in brca1 , all were defective for dna repair by either pathway . the hdr assay was accurate because all pathogenic mutants were defective for hdr , and all nonpathogenic variants were fully functional for hdr . repair by ssa accurately identified most variants , but several nonpathogenic variants were scored as defective or partially defective . these results indicated that specific amino acid residues of the brca1 protein have different effects in the two related dna repair pathways , and these results validate the hdr assay as highly correlative with brca1 - associated breast cancer . $F(s_{\text{adv}}, l =$"<multilabel>"$) = 1.0$ **Cos.** = 0.13 $MedSTS$ = 0.82	missense variants of uncertain clinical significance in the brca1 gene are a vexing problem in genetic counseling for women who have a family history of hereditary cancer . in this study , we evaluated the functions of 29 nonsynonymous polymorphisms of brca1 in two dna repair pathways . repair of double - strand breaks by homology - directed recombination (hdr) had been previously analyzed for 16 of these brca1 variants , and 13 more variants were analyzed in this study . all 29 variants were also analyzed for function in double - strand break repair by the single - strand annealing (ssa) pathway . we found that among the pathogenic mutations in brca1 , all were defective for dna repair by either pathway . the hdr assay was accurate because all missense mutants were defective for hdr , and all nonpathogenic variants were fully functional for hdr . repair by ssa accurately identified pathogenic variants , but several nonpathogenic variants were scored as defective or partially defective . these results indicated that specific amino acid residues of the brca1 protein have different effects in the two related dna repair pathways , and these results validate the hdr assay as highly correlative with of - and mutation studies . $F(s_{\text{adv}}, l =$"<multilabel>"$) = 1.0$ **Cos.** = 0.56 $MedSTS$ = 0.8	genetic variants of uncertain clinical significance in the brca1 gene are a vexing problem in genetic counseling for women who have a family history of breast cancer . in this study , we evaluated the functions of 29 missense variants of brca1 in two dna repair pathways . repair of double - strand breaks by homology - directed recombination (hdr) had been previously analyzed for 16 of these brca1 variants , and 13 more variants were analyzed in this study . all 29 variants were also analyzed for function in double - strand break repair by the single - strand annealing (ssa) pathway . we found that among the 28 variants in brca1 , all were defective for dna repair by either pathway . the hdr assay was accurate because all 15 mutants were defective for hdr , and all nonpathogenic variants were fully functional for hdr . repair by ssa accurately identified functional mutants , but several nonpathogenic mutations were scored as defective or partially defective . these results indicated that specific amino acid residues of the brca1 protein have different effects in the two related dna repair pathways , and these results validate the hdr assay as highly useful with single - associated breast cancer . $F(s_{\text{adv}}, l =$"<multilabel>"$) = 1.0$ **Cos.** = 0.73 $MedSTS$ = 0.81
Original	the punica granatum l . var . granatum (pomegranate) has been demonstrated to exert antitumor effects on various types of cancer cells . the present study aimed to evaluate the medicinal herbs punica granatum l . var . spinosa (apple punice) that are native to iran . this study was determined to test the possible cytotoxic activity and induction of apoptosis on human prostate cell lines . the effect of ethanol extracts of the herbs on the inhibition of cell proliferation was assessed by mtt colorimetric assay . pc3 cell lines treated with the extracts were analyzed for the induction of apoptosis by cell death detection (elisa) and tunel assay . dye exclusion analysis was performed for viability rate . our results demonstrated that the punica granatum l . var . spinosa extract dose dependently suppressed the proliferation of pc3 cells (ic (50) = 250 . 21 μg / ml) when compared with a chemotherapeutic anticancer drug (toxol) (vesper pharmaceuticals) with increased nucleosome production from apoptotic cells . the punica granatum l . var . spinosa extract attenuated the human prostate cell proliferation in vitro possibly by inducing apoptosis . the punica granatum l . var . spinosa is likely to be valuable for the treatment of some forms of human prostate cell line . $F(s, l =$"<multilabel>"$) = 1.0$	the punica granatum l . var . granatum (pomegranate) has been demonstrated to exert antitumor effects on various types of cancer cells . the present study aimed to evaluate the medicinal herbs punica granatum l . var . spinosa (apple punice) that are native to iran . this study was determined to test the possible cytotoxic activity and induction of apoptosis on human prostate cell lines . the effect of ethanol extracts of the herbs on the inhibition of cell proliferation was assessed by mtt colorimetric assay . pc3 cell lines treated with the extracts were analyzed for the induction of apoptosis by cell death detection (elisa) and tunel assay . dye exclusion analysis was performed for viability rate . our results demonstrated that the punica granatum l . var . spinosa extract dose dependently suppressed the proliferation of pc3 cells (ic (50) = 250 . 21 μg / ml) when compared with a chemotherapeutic anticancer drug (toxol) (vesper pharmaceuticals) with increased nucleosome production from apoptotic cells . the punica granatum l . var . spinosa extract attenuated the human prostate cell proliferation in vitro possibly by inducing apoptosis . the punica granatum l . var . spinosa is likely to be valuable for the treatment of some forms of human prostate cell line . $F(s, l =$"<multilabel>"$) = 1.0$	the punica granatum l . var . granatum (pomegranate) has been demonstrated to exert antitumor effects on various types of cancer cells . the present study aimed to evaluate the medicinal herbs punica granatum l . var . spinosa (apple punice) that are native to iran . this study was determined to test the possible cytotoxic activity and induction of apoptosis on human prostate cell lines . the effect of ethanol extracts of the herbs on the inhibition of cell proliferation was assessed by mtt colorimetric assay . pc3 cell lines treated with the extracts were analyzed for the induction of apoptosis by cell death detection (elisa) and tunel assay . dye exclusion analysis was performed for viability rate . our results demonstrated that the punica granatum l . var . spinosa extract dose dependently suppressed the proliferation of pc3 cells (ic (50) = 250 . 21 μg / ml) when compared with a chemotherapeutic anticancer drug (toxol) (vesper pharmaceuticals) with increased nucleosome production from apoptotic cells . the punica granatum l . var . spinosa extract attenuated the human prostate cell proliferation in vitro possibly by inducing apoptosis . the punica granatum l . var . spinosa is likely to be valuable for the treatment of some forms of human prostate cell line . $F(s, l =$"<multilabel>"$) = 1.0$

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	the punica granatum l . var . granatum (pomegranate) has been demonstrated to exert antitumor effects on various types of cancer cells . the present study aimed to evaluate the medicinal herbs punica granatum l . var . spinosa (apple punice) that are native to iran . this study was determined to test the possible antioxidant activity and induction of cytotoxicity on human pca cell lines . the effect of ethanol extracts of the herbs on the inhibition of cell proliferation was assessed by mtt colorimetric assay . pc3 cell lines treated with the extracts were analyzed for the induction of p53 by cell apoptosis detection (elisa) and tunel assay . dye exclusion analysis was performed for viability rate . our results demonstrated that the punica granatum l . var . spinosa extract dose dependently suppressed the proliferation of pc3 cells (ic (50) = 250 . 21 micrograms / ml) when compared with a chemotherapeutic model drug (toxol) (vesper pharmaceuticals) with increased nucleosome number from apoptotic cells . the punica granatum l . var . spinosa extract attenuated the human pc cell proliferation in vitro possibly by inducing pge2 . the punica granatum l . var . spinosa is likely to be valuable for the treatment of some forms of human pc cell line . $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 1.0 **Cos.** = 0.32 $MedSTS$ = 0.93	the punica granatum l . var . granatum (pomegranate) has been demonstrated to exert antitumor effects on various types of cancer cells . the present study aimed to evaluate the medicinal herbs punica granatum l . var . spinosa (apple punice) that are native to iran . this study was determined to test the possible antioxidant activity and induction of cytotoxicity on human pc3 cell lines . the effect of ethanol extracts of the leaves on the inhibition of cell proliferation was assessed by mtt colorimetric assay . pc3 cell lines treated with the extracts were analyzed for the induction of caspases by cell death detection (elisa) and ldh assay . dye viability analysis was performed for viability rate . our results demonstrated that the punica granatum l . var . spinosa extract dose dependently suppressed the proliferation of pc3 cells (ic (50) = 250 . 21 μg / ml) when compared with a chemotherapeutic anticancer drug (toxol) (vesper pharmaceuticals) with increased apoptosis death from apoptotic cells . the punica granatum l . var . spinosa extract attenuated the human prostate cell proliferation in vitro possibly by inducing p53 . the punica granatum l . var . spinosa is likely to be valuable for the treatment of some forms of human prostatic cell line . $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 1.0 **Cos.** = 0.44 $MedSTS$ = 0.84	the punica granatum l . var . granatum (pomegranate) has been demonstrated to exert antitumor effects on various types of cancer cells . the present study aimed to evaluate the medicinal properties punica granatum l . var . spinosa (citrus punice) that are native to india . this study was determined to test the possible cytotoxic activity and induction of differentiation on human tumor cell lines . the effect of ethanol extracts of the fruit on the inhibition of cell proliferation was assessed by mtt colorimetric assay . pc3 cell lines treated with the extracts were analyzed for the induction of apoptosis by cell death detection (elisa) and immunoblot assay . dye exclusion analysis was performed for viability rate . our results demonstrated that the punica granatum l . var . spinosa extract dose dependently suppressed the proliferation of pc3 cells (ic (50) = 250 . 21 μg / ml) when compared with a chemotherapeutic anticancer drug (toxol) (vesper pharmaceuticals) with increased apoptosis death from apoptotic cells . the punica granatum l . var . spinosa extract attenuated the human pc3 cell proliferation in vitro possibly by inducing apoptosis . the punica granatum l . var . spinosa is likely to be valuable for the treatment of some forms of human pc3 cell line . $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 1.0 **Cos.** = 0.54 $MedSTS$ = 0.8
Original	objective although downregulation of neural cell adhesion molecule (ncam) has been correlated with poor prognosis in colorectal cancer (crc) , it is also possible that colon cancer spreading comes from reducing tumor cell adhesion through ncam polysialylation , as occurs in lung carcinoma or wilms ' tumor . methods to prove this hypothesis , we have performed a prospective study on tumor and control specimens from 39 crc patients , which were immunostained for ncam and psa (polysialic acid) expression . results tumor versus control expression of ncam and psa epitopes in tissue specimens , as well as correlation between tumor expression and clinicopathological features , were statistically analyzed . results showed a low constitutive expression of ncam and psa (psa - ncam) in control tissue , which reached a statistically significant increase in the tumor tissue . likewise , the presence and number of lymph node metastases at surgery were correlated with ncam expression and psa / ncam coexpression . conclusions these data highlight the importance of taking into account psa - associated epitopes when dealing with ncam cell expression studies in tumor development and progression . the analysis of psa and ncam expression in crc suggests a new way , other than downregulation of ncam , in order to escape contact inhibition and promote cell tumor spreading in colorectal cancer . $F(\boldsymbol{s}, l =$"<multilabel>") = 1.0	objective although downregulation of neural cell adhesion molecule (ncam) has been correlated with poor prognosis in colorectal cancer (crc) , it is also possible that colon cancer spreading comes from reducing tumor cell adhesion through ncam polysialylation , as occurs in lung carcinoma or wilms ' tumor . methods to prove this hypothesis , we have performed a prospective study on tumor and control specimens from 39 crc patients , which were immunostained for ncam and psa (polysialic acid) expression . results tumor versus control expression of ncam and psa epitopes in tissue specimens , as well as correlation between tumor expression and clinicopathological features , were statistically analyzed . results showed a low constitutive expression of ncam and psa (psa - ncam) in control tissue , which reached a statistically significant increase in the tumor tissue . likewise , the presence and number of lymph node metastases at surgery were correlated with ncam expression and psa / ncam coexpression . conclusions these data highlight the importance of taking into account psa - associated epitopes when dealing with ncam cell expression studies in tumor development and progression . the analysis of psa and ncam expression in crc suggests a new way , other than downregulation of ncam , in order to escape contact inhibition and promote cell tumor spreading in colorectal cancer . $F(\boldsymbol{s}, l =$"<multilabel>") = 1.0	objective although downregulation of neural cell adhesion molecule (ncam) has been correlated with poor prognosis in colorectal cancer (crc) , it is also possible that colon cancer spreading comes from reducing tumor cell adhesion through ncam polysialylation , as occurs in lung carcinoma or wilms ' tumor . methods to prove this hypothesis , we have performed a prospective study on tumor and control specimens from 39 crc patients , which were immunostained for ncam and psa (polysialic acid) expression . results tumor versus control expression of ncam and psa epitopes in tissue specimens , as well as correlation between tumor expression and clinicopathological features , were statistically analyzed . results showed a low constitutive expression of ncam and psa (psa - ncam) in control tissue , which reached a statistically significant increase in the tumor tissue . likewise , the presence and number of lymph node metastases at surgery were correlated with ncam expression and psa / ncam coexpression . conclusions these data highlight the importance of taking into account psa - associated epitopes when dealing with ncam cell expression studies in tumor development and progression . the analysis of psa and ncam expression in crc suggests a new way , other than downregulation of ncam , in order to escape contact inhibition and promote cell tumor spreading in colorectal cancer . $F(\boldsymbol{s}, l =$"<multilabel>") = 0.99

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	objective although downregulation of neuronal neural cell molecule (ncam) has been correlated with poor prognosis in colorectal cancer (crc) , it is also possible that colon cancer aggressiveness comes from reducing tumor cell adhesion through ncam polysialylation , as occurs in lung carcinoma or wilms ' tumor . methods to prove this hypothesis , we have performed a prospective study on tumor and control specimens from 39 crc patients , which were immunostained for ncam and pa (polysialic acid) expression . results tumor versus control expression of ncam and psa epitopes in tissue specimens , as well as correlation between tumor expression and clinicopathological features , were statistically analyzed . results showed a low constitutive expression of ncam and psa (anti - ncam) in control tissue , which reached a statistically significant increase in the tumor tissue . likewise , the presence and number of regional node metastases at surgery were correlated with ncam expression and psa / ncam coexpression . conclusions these data highlight the importance of taking into account psa - associated epitopes when dealing with ncam cell expression studies in tumor development and progression . the analysis of psa and ncam expression in crc suggests a new way , other than downregulation of ncam , in order to escape nc inhibition and thus cell cell spread in colorectal cancer . $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 1.0 **Cos.** = -0.16 $MedSTS = 0.74$	objective although downregulation of neuronal cell adhesion molecule (ncam) has been correlated with poor prognosis in colorectal cancer (crc) , it is also possible that colon cancer spreading comes from aberrant tumor cell adhesion through nc adhesion , as occurs in lung carcinoma or wilms ' tumor . methods to prove this hypothesis , we have performed a prospective study on tumor and control specimens from 39 crc patients , which were immunostained for ncam and psma (polysialic acid) expression . results tumor versus control expression of ncam and specific epitopes in tissue specimens , as well as correlation between tumor expression and clinicopathological features , were statistically analyzed . results showed a low constitutive expression of ncam and psa (serine - ncam) in control tissue , which reached a statistically significant increase in the tumor tissue . likewise , the presence and number of lymph node metastases at surgery were correlated with ncam expression and psa / ncam coexpression . conclusions these data highlight the importance of taking into account cell - associated epitopes when dealing with ncam cell expression studies in tumor development and progression . the analysis of pca and ncam expression in crc suggests a new way , other than downregulation of ncam , in order to escape the metastasis and promote cell tumor spreading in colorectal cancer . $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 0.98 **Cos.** = 0.19 $MedSTS = 0.7$	objective although downregulation of cell - activation molecule (psa) has been correlated with poor prognosis in colorectal cancer (crc) , it is also possible that colon cancer metastasis comes from reducing tumor cell proliferation through antigen upregulation , as occurs in lung carcinoma or wilms ' tumor . methods to prove this hypothesis , we have performed a prospective study on tumor and control specimens from 39 crc patients , which were immunostained for ncam and protein (polysialic acid) expression . results tumor versus control expression of ncam and psa epitopes in tissue specimens , as well as correlation between tumor expression and clinicopathological features , were statistically analyzed . results showed a low constitutive expression of ncam and psa (psa - ncam) in control tissue , which reached a statistically significant increase in the tumor tissue . likewise , the presence and number of lymph node metastases at surgery were correlated with ncam expression and psa / ncam coexpression . conclusions these data highlight the importance of taking into account psa - associated epitopes when dealing with ncam cell expression studies in tumor development and progression . the analysis of cd44 and ncam expression in crc suggests a new way , other than downregulation of ncam , in order to escape contact inhibition and promote cell tumor growth in colorectal cancer . $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 0.99 **Cos.** = 0.37 $MedSTS = 0.7$
Original	objective to analyze histological factors not routinely assessed as potential prognostic factors in renal cell carcinoma , such as tumor necrosis , microscopic vascular invasion , and sinus fat invasion . materials and methods a retrospective , analytical study was conducted of surgical specimens from 139 patients with localized renal cell carcinoma who underwent nephrectomy from 1993 to 2005 . tumor necrosis , microscopic vascular invasion , and sinus fat invasion were analyzed and compared to the classical factors : tnm classification , fuhrman grade , and tumor size . for statistical analysis , variables analyzed were categorized as pt1 , 2 vs pt3 , 4 ; fuhrman grade 1 , 2 vs 3 , 4 ; tumor size < 7 cm vs > or = 7cm ; tumor necrosis vs no tumor necrosis ; microvascular invasion of sinus fat vs no invasion . cancer - specific survival probability and disease - free survival were calculated . a descriptive and analytical statistical analysis was performed using logistic regression for univariate and multivariate analyses . dependent variables were used to analyze cancer - specific survival rates . disease - free survival was estimated using a cox regression model and kaplan - meier curves . results in the univariate analysis , all variables analyzed had a significant influence on death for renal cell carcinoma . in the multivariate analysis , the variable having the greatest influence was fuhrman grade (p = 0 , 032) . $F(\boldsymbol{s}, l =$"<multilabel>") = 1.0	objective to analyze histological factors not routinely assessed as potential prognostic factors in renal cell carcinoma , such as tumor necrosis , microscopic vascular invasion , and sinus fat invasion . materials and methods a retrospective , analytical study was conducted of surgical specimens from 139 patients with localized renal cell carcinoma who underwent nephrectomy from 1993 to 2005 . tumor necrosis , microscopic vascular invasion , and sinus fat invasion were analyzed and compared to the classical factors : tnm classification , fuhrman grade , and tumor size . for statistical analysis , variables analyzed were categorized as pt1 , 2 vs pt3 , 4 ; fuhrman grade 1 , 2 vs 3 , 4 ; tumor size < 7 cm vs > or = 7cm ; tumor necrosis vs no tumor necrosis ; microvascular invasion of sinus fat vs no invasion . cancer - specific survival probability and disease - free survival were calculated . a descriptive and analytical statistical analysis was performed using logistic regression for univariate and multivariate analyses . dependent variables were used to analyze cancer - specific survival rates . disease - free survival was estimated using a cox regression model and kaplan - meier curves . results in the univariate analysis , all variables analyzed had a significant influence on death for renal cell carcinoma . in the multivariate analysis , the variable having the greatest influence was fuhrman grade (p = 0 , 032) . $F(\boldsymbol{s}, l =$"<multilabel>") = 1.0	objective to analyze histological factors not routinely assessed as potential prognostic factors in renal cell carcinoma , such as tumor necrosis , microscopic vascular invasion , and sinus fat invasion . materials and methods a retrospective , analytical study was conducted of surgical specimens from 139 patients with localized renal cell carcinoma who underwent nephrectomy from 1993 to 2005 . tumor necrosis , microscopic vascular invasion , and sinus fat invasion were analyzed and compared to the classical factors : tnm classification , fuhrman grade , and tumor size . for statistical analysis , variables analyzed were categorized as pt1 , 2 vs pt3 , 4 ; fuhrman grade 1 , 2 vs 3 , 4 ; tumor size < 7 cm vs > or = 7cm ; tumor necrosis vs no tumor necrosis ; microvascular invasion of sinus fat vs no invasion . cancer - specific survival probability and disease - free survival were calculated . a descriptive and analytical statistical analysis was performed using logistic regression for univariate and multivariate analyses . dependent variables were used to analyze cancer - specific survival rates . disease - free survival was estimated using a cox regression model and kaplan - meier curves . results in the univariate analysis , all variables analyzed had a significant influence on death for renal cell carcinoma . in the multivariate analysis , the variable having the greatest influence was fuhrman grade (p = 0 , 032) . $F(\boldsymbol{s}, l =$"<multilabel>") = 0.93

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	objective to analyze clinicopathological factors not routinely assessed as potential prognostic factors in clear cell carcinoma , such as tumor necrosis , microscopic venous permeation , and sinus fat involvement . materials and methods a retrospective , analytical study was conducted of surgical specimens from 139 patients with localized renal cell carcinoma who underwent nephrectomy from 1993 to 2005 . tumor invasion , microscopic tumor invasion , and sinus fat invasion were analyzed and compared to the classical factors : tnm classification , fuhrman grade , and tumor size . for statistical analysis , variables analyzed were categorized as pt1 , 2 vs pt3 , 4 ; fuhrman grade 1 , 2 vs 3 , 4 ; tumor size < 7 cm vs > or = 7cm ; tumor thrombus vs no microscopic invasion ; invasion presence of sinus fat vs no invasion . cancer - specific survival probability and disease - free survival were calculated . a descriptive and analytical statistical analysis was performed using logistic regression for univariate and multivariate analyses . dependent variables were used to analyze cancer - specific survival rates . disease - free survival was estimated using a cox regression model and kaplan - meier curves . results in the univariate analysis , all variables analyzed had a significant influence on death for renal cell carcinoma . in the multivariate analysis , the variable having the greatest influence was fuhrman grade (p = 0 , 032) . $F(s_{\text{adv}}, l =$ “<multilabel>”$) = 0.99$ **Cos.** = 0.09 $MedSTS$ = 0.66	objective to analyze several factors not routinely assessed as potential prognostic factors in renal cell carcinoma , such as tumor necrosis , microscopic vascular invasion , and lymph fat invasion . materials and methods a retrospective , analytical study was conducted of surgical specimens from 139 patients with localized renal cell carcinoma who underwent nephrectomy from 1993 to 2005 . tumor invasion , microscopic vascular permeation , and microscopic vascular invasion were analyzed and compared to the classical factors : histological classification , fuhrman grade , and tumor size . for statistical analysis , variables analyzed were categorized as pt1 , 2 vs pt3 , 4 ; fuhrman grade 1 , 2 vs 3 , 4 ; tumor size < 7 cm vs > or = 7cm ; tumor infiltration vs no tumor permeation ; invasion lymph of vascular fat vs no invasion . cancer - specific survival probability and disease - free survival were calculated . a descriptive and analytical statistical analysis was performed using logistic regression for univariate and multivariate analyses . dependent variables were used to analyze cancer - specific survival rates . disease - free survival was estimated using a cox regression model and kaplan - meier curves . results in the univariate analysis , all variables analyzed had a significant influence on death for renal cell carcinoma . in the multivariate analysis , the variable having the greatest influence was fuhrman grade (p = 0 , 032) . $F(s_{\text{adv}}, l =$ “<multilabel>”$) = 0.97$ **Cos.** = 0.28 $MedSTS$ = 0.73	objective to analyze histological factors not routinely assessed as potential prognostic factors in renal cell carcinoma , such as tumor necrosis , microscopic vascular invasion , and lymph lymphatic invasion . materials and methods a retrospective , analytical study was conducted of surgical specimens from 139 patients with localized renal cell carcinoma who underwent lymphadenectomy from 1993 to 2005 . tumor metastasis , microscopic vascular invasion , and lymph fat metastases were analyzed and compared to the classical factors : histological classification , fuhrman grade , and tumor size . for statistical analysis , variables analyzed were categorized as pt1 , 2 vs pt3 , 4 ; fuhrman grade 1 , 2 vs 3 , 4 ; tumor size < 7 cm vs > or = 7cm ; tumor necrosis vs no tumor necrosis ; neph lymph of vascular necrosis vs no necrosis . cancer - specific survival probability and disease - free survival were calculated . a descriptive and analytical statistical analysis was performed using logistic regression for univariate and multivariate analyses . dependent variables were used to analyze cancer - specific survival rates . disease - free survival was estimated using a cox regression model and kaplan - meier curves . results in the univariate analysis , all variables analyzed had a significant influence on death for renal cell carcinoma . in the multivariate analysis , the variable having the greatest influence was fuhrman grade (p = 0 , 032) . $F(s_{\text{adv}}, l =$ “<multilabel>”$) = 0.96$ **Cos.** = 0.62 $MedSTS$ = 0.67
Original	dna - protein cross - links (dpcs) are formed upon exposure to a variety of chemical and physical agents and pose a threat to genomic integrity . in particular , acrolein and related aldehydes produce dpcs , although the chemical linkages for such cross - links have not been identified . here , we report that oligodeoxynucleotides containing 1 , n (2) - deoxyguanosine adducts of acrolein , crotonaldehyde , and trans - 4 - hydroxynonenal can form cross - links with the tetrapeptide lys - trp - lys - lys . we concluded that complex formation is mediated by a schiff base linkage because dna - peptide complexes were covalently trapped following reduction with sodium cyanoborohydride , and pre - reduction of adducted dnas inhibited complex formation . a previous nmr study demonstrated that duplex dna catalyzes ring opening for the acrolein - derived gamma - hydroxy - 1 , n (2) - propanodeoxyguanosine adduct to yield an aldehydic function (de los santos , c . , zaliznyak , t . , and johnson , f . (2001) j . biol . chem . 276 , 9077 - 9082) . consistent with this earlier observation , the adducts under investigation were more reactive in duplex dna than $F(s, l =$ “<multilabel>”$) = 1.0$	dna - protein cross - links (dpcs) are formed upon exposure to a variety of chemical and physical agents and pose a threat to genomic integrity . in particular , acrolein and related aldehydes produce dpcs , although the chemical linkages for such cross - links have not been identified . here , we report that oligodeoxynucleotides containing 1 , n (2) - deoxyguanosine adducts of acrolein , crotonaldehyde , and trans - 4 - hydroxynonenal can form cross - links with the tetrapeptide lys - trp - lys - lys . we concluded that complex formation is mediated by a schiff base linkage because dna - peptide complexes were covalently trapped following reduction with sodium cyanoborohydride , and pre - reduction of adducted dnas inhibited complex formation . a previous nmr study demonstrated that duplex dna catalyzes ring opening for the acrolein - derived gamma - hydroxy - 1 , n (2) - propanodeoxyguanosine adduct to yield an aldehydic function (de los santos , c . , zaliznyak , t . , and johnson , f . (2001) j . biol . chem . 276 , 9077 - 9082) . consistent with this earlier observation , the adducts under investigation were more reactive in duplex dna than $F(s, l =$ “<multilabel>”$) = 1.0$	dna - protein cross - links (dpcs) are formed upon exposure to a variety of chemical and physical agents and pose a threat to genomic integrity . in particular , acrolein and related aldehydes produce dpcs , although the chemical linkages for such cross - links have not been identified . here , we report that oligodeoxynucleotides containing 1 , n (2) - deoxyguanosine adducts of acrolein , crotonaldehyde , and trans - 4 - hydroxynonenal can form cross - links with the tetrapeptide lys - trp - lys - lys . we concluded that complex formation is mediated by a schiff base linkage because dna - peptide complexes were covalently trapped following reduction with sodium cyanoborohydride , and pre - reduction of adducted dnas inhibited complex formation . a previous nmr study demonstrated that duplex dna catalyzes ring opening for the acrolein - derived gamma - hydroxy - 1 , n (2) - propanodeoxyguanosine adduct to yield an aldehydic function (de los santos , c . , zaliznyak , t . , and johnson , f . (2001) j . biol . chem . 276 , 9077 - 9082) . consistent with this earlier observation , the adducts under investigation were more reactive in duplex dna than $F(s, l =$ “<multilabel>”$) = 1.0$

	Vanilla	Adversarial	FAR-IG
Adversarial	dna - protein cross - links (dpcs) are formed upon exposure to a variety of chemical and physical agents and pose a threat to genomic integrity . in particular , acrolein and related molecules produce dpcs , although the chemical linkages for such cross - links have not been identified . here , we report that oligodeoxynucleotides containing 1 , n (2) - diol esters of acrolein , crotonaldehyde , and trans - 4 - hydroxynonenal can form cross - linking with the tetrapeptide lys - trp - lys - lys . we concluded that complex formation is mediated by a schiff base linkage because dna - peptide complexes were selectively trapped following reduction with sodium cyanoborohydride , and pre - reduction of adducted dnas inhibited complex formation . a previous nmr study demonstrated that duplex dna catalyzes ring opening for the acrolein - derived gamma - hydroxy - 1 , n (2) - propanodeoxyguanosine radical to yield an aldehydic function (de los santos , c . , zaliznyak , t . , and johnson , f . (2001) j . biol . chem . 276 , 9077 - 9082) . consistent with this earlier observation , the peptides under investigation were more efficiently in this buffer than $F(\boldsymbol{s}_{\text{adv}},\ l =$ "<multilabel>") = 1.0 **Cos.** = 0.33 $MedSTS$ = 0.62	dna - protein cross - links (dpcs) are formed upon exposure to a variety of chemical and physical agents and pose a threat to genomic integrity . in particular , acrolein and related molecules produce dpcs , although the chemical linkages for such cross - links have not been identified . here , we report that oligodeoxynucleotides containing 1 , n (2) - deoxyguanosine analogues of acrolein , crotonaldehyde , and trans - 4 - hydroxynonenal can form cross - linkages with the sequences lys - trp - lys - lys . we concluded that complex formation is mediated by a single base linkage because dna - peptide complexes were not trapped following reduction with sodium azide , and pre - reduction of adducted dnas inhibited complex formation . a previous nmr study demonstrated that duplex dna catalyzes the opening for the acrolein - derived gamma - hydroxy - 1 , n (2) - propanodeoxyguanosine molecule to yield an aldehydic function (de los santos , c . , zaliznyak , t . , and johnson , f . (2001) j . biol . chem . 276 , 9077 - 9082) . consistent with this earlier observation , the bases under investigation were more reactive in duplex dna than $F(\boldsymbol{s}_{\text{adv}},\ l =$ "<multilabel>") = 1.0 **Cos.** = 0.54 $MedSTS$ = 0.64	dna - protein cross - links (dpcs) are formed upon exposure to a variety of chemical and physical agents and pose a threat to genomic integrity . in particular , acrolein and related adducts produce dpcs , although the chemical linkages for such cross - links have not been identified . here , we report that oligodeoxynucleotides containing 1 , n (2) - deoxyguanosine conjugates of acrolein , crotonaldehyde , and trans - 4 - hydroxynonenal can form cross - links with the sequences lys - lys - lys - lys . we concluded that dpc formation is mediated by a dna base linkage because dna - adduct sites were covalently trapped following reduction with sodium cyanoborohydride , and pre - reduction of adducted dnas inhibited complex formation . a previous nmr study demonstrated that duplex dna catalyzes ring opening for the acrolein - derived gamma - keto - 1 , n (2) - propanodeoxyguanosine adduct to yield an aldehydic function (de los santos , c . , zaliznyak , t . , and johnson , f . (2001) j . biol . chem . 276 , 9077 - 9082) . consistent with this earlier observation , the linkages under investigation were more reactive in duplex dna than $F(\boldsymbol{s}_{\text{adv}},\ l =$ "<multilabel>") = 1.0 **Cos.** = 0.66 $MedSTS$ = 0.67
Original	verbascum thapsus commonly known as ' mullein ' is part of a large family of scrophulariaceae consisting of more than 360 species . from antiquity verbascum thapsus has been used as a medicinal herb , it contains diverse polysaccharides , iroid glycosides , flavonoids , saponins , volatile oils and phenylentanoids . inducible nitric oxide synthase (inos) represents one of the three isoforms that produce nitric oxide using l - arginine as a substrate in response to an increase in superoxide anion activated by nf - kb . it is implicated in different pathophysiological events and its expression increases greatly during an inflammatory process , due to oxidative stress and the activation of the enzymes of the antioxidant network such as sod , cat and gpx . in this study an inflammatory state was reproduced by treating thp - 1 cells (human myelomonocytic leukaemia) with pro - inflammatory stimuli , such as lps and ifn - gamma , obtaining an up - regulation both in the expression and in the activity of inos . the aim of the work was to investigate the antiinflammatory action of verbascoside using a concentration of 100 mum . the results show a significant decrease of the expression and activity of inos , extracellular o (2) (-) production , sod , cat and gpx activity when the cells were treated $F(\boldsymbol{s},\ l =$ "<multilabel>") = 1.0	verbascum thapsus commonly known as ' mullein ' is part of a large family of scrophulariaceae consisting of more than 360 species . from antiquity verbascum thapsus has been used as a medicinal herb , it contains diverse polysaccharides , iroid glycosides , flavonoids , saponins , volatile oils and phenylentanoids . inducible nitric oxide synthase (inos) represents one of the three isoforms that produce nitric oxide using l - arginine as a substrate in response to an increase in superoxide anion activated by nf - kb . it is implicated in different pathophysiological events and its expression increases greatly during an inflammatory process , due to oxidative stress and the activation of the enzymes of the antioxidant network such as sod , cat and gpx . in this study an inflammatory state was reproduced by treating thp - 1 cells (human myelomonocytic leukaemia) with pro - inflammatory stimuli , such as lps and ifn - gamma , obtaining an up - regulation both in the expression and in the activity of inos . the aim of the work was to investigate the antiinflammatory action of verbascoside using a concentration of 100 mum . the results show a significant decrease of the expression and activity of inos , extracellular o (2) (-) production , sod , cat and gpx activity when the cells were treated $F(\boldsymbol{s},\ l =$ "<multilabel>") = 1.0	verbascum thapsus commonly known as ' mullein ' is part of a large family of scrophulariaceae consisting of more than 360 species . from antiquity verbascum thapsus has been used as a medicinal herb , it contains diverse polysaccharides , iroid glycosides , flavonoids , saponins , volatile oils and phenylentanoids . inducible nitric oxide synthase (inos) represents one of the three isoforms that produce nitric oxide using l - arginine as a substrate in response to an increase in superoxide anion activated by nf - kb . it is implicated in different pathophysiological events and its expression increases greatly during an inflammatory process , due to oxidative stress and the activation of the enzymes of the antioxidant network such as sod , cat and gpx . in this study an inflammatory state was reproduced by treating thp - 1 cells (human myelomonocytic leukaemia) with pro - inflammatory stimuli , such as lps and ifn - gamma , obtaining an up - regulation both in the expression and in the activity of inos . the aim of the work was to investigate the antiinflammatory action of verbascoside using a concentration of 100 mum . the results show a significant decrease of the expression and activity of inos , extracellular o (2) (-) production , sod , cat and gpx activity when the cells were treated $F(\boldsymbol{s},\ l =$ "<multilabel>") = 1.0

	VANILLA	ADVERSARIAL	FAR-IG
Adversarial	verbascum thapsus commonly known as ' mullein ' is part of a large family of scrophulariaceae consisting of more than 360 species . from antiquity verbascum thapsus has been used as a medicinal herb , it contains diverse polysaccharides , iroid glycosides , flavonoids , saponins , volatile oils and phenylentanoids . inducible nitric oxide synthase (its) represents one of the three isoforms that produce nitric oxide using l - arginine as a substrate in response to an increase in radical anion activated by nf - kb . it is implicated in different pathophysiological events and its expression increases greatly during an activation process , due to redox stress and the activation of the enzymes of the antioxidant network such as sod , cat and gpx . in this study an activation model was reproduced by treating l - 1 cells (human myelomonocytic leukaemia) with pro - oxidant stimuli , such as pma and ifn - gamma , obtaining an up - regulation both in the expression and in the activity of inos . the aim of the work was to investigate the inhibitory action of verbascoside using a concentration of 100 mum . the results show a significant decrease of the expression and activity of inos , extracellular o (2) (-) production , sod , cat and gpx activity when the stimulation were treated $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 0.9 **Cos.** = -0.01 $MedSTS$ = 0.62	verbascum thapsus commonly known as ' mullein ' is part of a large family of scrophulariaceae consisting of more than 360 species . from antiquity verbascum thapsus has been used as a medicinal herb , it contains diverse polysaccharides , iroid glycosides , flavonoids , saponins , volatile oils and phenylentanoids . inducible no oxide synthase (enos) represents one of the three isoforms that produce adenosine oxide using l - tyrosine as a substrate in response to an increase in inflammatory , activated by nf - kb . it is implicated in different pathophysiological events and its expression increases greatly during an inflammatory process , due to oxidative stress and the activation of the enzymes of the antioxidant network such as sod , cat and gpx . in this study an inflamed state was reproduced by treating thp - 1 cells (human myelomonocytic leukaemia) with pro - inflammatory stimuli , such as lps and ifn - gamma , obtaining an up - regulation both in the expression and in the activity of is . the aim of the work was to investigate the antioxidant action of verbascoside using a concentration of 100 mum . the results show a significant decrease of the expression and activity of nnos , extracellular no (2) (-) production , sod , cat and gpx activity when the cells were treated $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 0.98 **Cos.** = 0.4 $MedSTS$ = 0.62	verbascum thapsus commonly known as ' mullein ' is part of a large family of scrophulariaceae consisting of more than 360 species . from antiquity verbascum thapsus has been used as a medicinal herb , it contains diverse polysaccharides , iroid glycosides , flavonoids , saponins , volatile oils and phenylentanoids . inducible inducible monoxide synthase (ias) represents one of the three isoforms that produce nitric oxide using l - tyrosine as a substrate in response to an increase in hydroxyl anion activated by nf - kb . it is implicated in different pathophysiological events and its expression increases greatly during an inflammatory process , due to oxidant stress and the activation of the enzymes of the oxidative network such as sod , cat and gpx . in this study an activation state was reproduced by treating thp - 1 cells (human myelomonocytic leukaemia) with pro - inflammatory stimuli , such as lps and ifn - gamma , obtaining an up - regulation both in the expression and in the activity of enos . the aim of the work was to investigate the antioxidant action of verbascoside using a concentration of 100 mum . the results show a significant decrease of the expression and activity of nnos , extracellular o (2) (-) production , sod , cat and gpx activity when the cells were treated $F(\boldsymbol{s}_{\text{adv}}, l =$"<multilabel>") = 0.99 **Cos.** = 0.53 $MedSTS$ = 0.62