

# Evidence annotation guide

## 1 Task introduction

In this task, you will verify or refute scientific claims using evidence from the abstracts of a scientific article. For more detail on claims, see Section 2. Figure 1 shows the annotation interface. Follow these steps to verify a claim:

The screenshot displays the annotation interface for a claim. On the left, a 'Claim' box contains the text: 'In breast cancer, the loss of myoepithelial cells promotes the transition of ductal carcinoma in situ to invasive carcinoma.' Below this, the title 'Regulation of in situ to invasive breast carcinoma transition.' is shown, followed by an abstract snippet. The abstract snippet includes two numbered sentences: '1 The transition of ductal carcinoma in situ (DCIS) to invasive carcinoma is a poorly understood key event in breast tumor progression. Here, we analyzed the role of myoepithelial cells and fibroblasts in the progression of in situ carcinomas using a model of human DCIS and primary breast tumors. Progression to invasion was promoted by fibroblasts and inhibited by normal myoepithelial cells. 4 Molecular profiles of isolated luminal epithelial and myoepithelial cells identified an intricate interaction network involving TGFbeta, Hedgehog, cell adhesion, and p63 required for myoepithelial cell differentiation, the elimination of which resulted in loss of myoepithelial cells and progression to invasion.' On the right, a 'Submission 1' box shows the date 'Thu Mar 5 at 02:10 PM (18 days ago)'. Below this, a 'New Submission' section contains three buttons: 'ADD EVIDENCE', 'SKIP', and 'SUBMIT (1)'. A toggle switch for 'Flag submission as need attention' is also present. The 'Evidence Set' section shows two selected pieces of evidence. The first is labeled '1' and is marked as 'Supplemental'. The second is labeled '4' and is marked as 'Primary'. Below the evidence sets, a 'Level of Support / Contradict' section contains four buttons: 'FULL SUPPORT', 'PART SUPPORT', 'PART CONTRADICT', and 'FULL CONTRADICT'.

Figure 1: The annotation interface, with evidence selected.

1. Read the claim in the “claim” box.
2. If you spend more than a minute trying to understand the claim and can’t figure out what’s going on, flag the claim for further review by flipping the “Flag submission as need attention” switch at the top right of the interface, and write a note explaining why you skipped the claim. Then hit the “Skip” button to go on to the next example. Otherwise, go on to the next step.
3. Below the claim box, you will see the title and abstract of a scientific paper.
4. Identify all evidence sets in the abstract, as defined in Section 3. Sometimes, the abstract will not provide any relevant evidence, as in Figure 2. When this occurs, hit the “Submit” button with no evidence sets selected. *Do not* hit the “Skip” button.
5. If you have any concerns about the claim or the evidence you selected, flag the claim for further review and write a note explaining why you flagged the submission. Some claims are just really tough. When in doubt, flag it.

Figure 2: The abstract does not mention LRBA at all. It provides no evidence.

## 2 Claims

Each claim should be a **single sentence expressing a finding about one aspect of a scientific entity or process**. Figure 3 shows a “good” claim, along with two “bad” ones:

- Claim 1 is a “good” claim since it states a single finding (lung cancer) about a single scientific process (smoking).
- Claim 2 is “bad” since it states two separate findings (lung cancer and discolored teeth) about a single process (smoking).
- Claim 3 is “bad” since it states a finding (lung cancer) about two different scientific processes (smoking and asbestos exposure).

• <b>Claim 1 (Good):</b> Smoking causes lung cancer.
• <b>Claim 2 (Bad):</b> Smoking causes lung cancer and discolored teeth.
• <b>Claim 3 (Bad):</b> Smoking and asbestos exposure both cause lung cancer.

Figure 3: One good claim and two bad ones. The bad ones are harder to fact-check and should be flagged.

Claim 1 is easier to fact-check than the other two . If you’re shown a “bad” claim during the annotation process, do your best to find evidence. In addition, flag the claim and write a note explaining the issue. A real example of a “bad” claim is shown Figure 14, in Section 3.3.2.

### 3 Evidence sets

An **evidence set** is a collection of sentences from the abstract that provide support or contradiction for the given claim. To decide whether a collection of sentences is an evidence set, ask yourself, “If I were shown only these sentences, could I reasonably conclude that the claim is true (or false)”?

- Evidence sets should be *minimal*. If you can remove a sentence from the evidence set and the remaining sentences are sufficient for support / contradiction, you should remove it.
- There may be multiple evidence sets in a given abstract.

#### 3.1 Types of evidence sentences

Evidence sets are composed of two types of sentences:

- **Primary sentences** state a fact or finding that verifies or contradicts the claim. You can only use a primary sentence in a **single evidence set**. The primary sentences in a single evidence set should be *minimal*. If you can remove a sentence and the remaining sentences still verify the claim, remove it. If the removed sentence can be used as a primary sentence itself, create a new evidence set using it.
  - Primary sentences can state new findings reported in the abstract, or findings from previous papers, as shown in Figure 4
  - An annotator wrote: “Any definitive sentence from the paragraph that I could point to and say the claim is true should be labeled a primary sentence. This is regardless of whether it reports a result from the study or not”.

– <b>claim:</b> Aspirin reduces headaches
– <b>Evidence 1 (primary):</b> Previous work showed that aspirin reduces headaches.
– <b>Evidence 2 (primary):</b> Our results show that aspirin reduces headaches.

Figure 4: Both these evidence sentences are *primary*. Each sentence provides convincing evidence that the claim is true. In particular, Evidence 1 is primary, even though it reports a finding that was established in a different paper.

- **Supplemental sentences** provide supplemental information that is necessary to establish the context or setting for the primary sentences — for instance, the intervention, population under study, cell line, or species. If multiple sentences could be used to provide supplemental information, pick the earliest one. You can use supplemental sentences in **as many evidence sets as you want**.

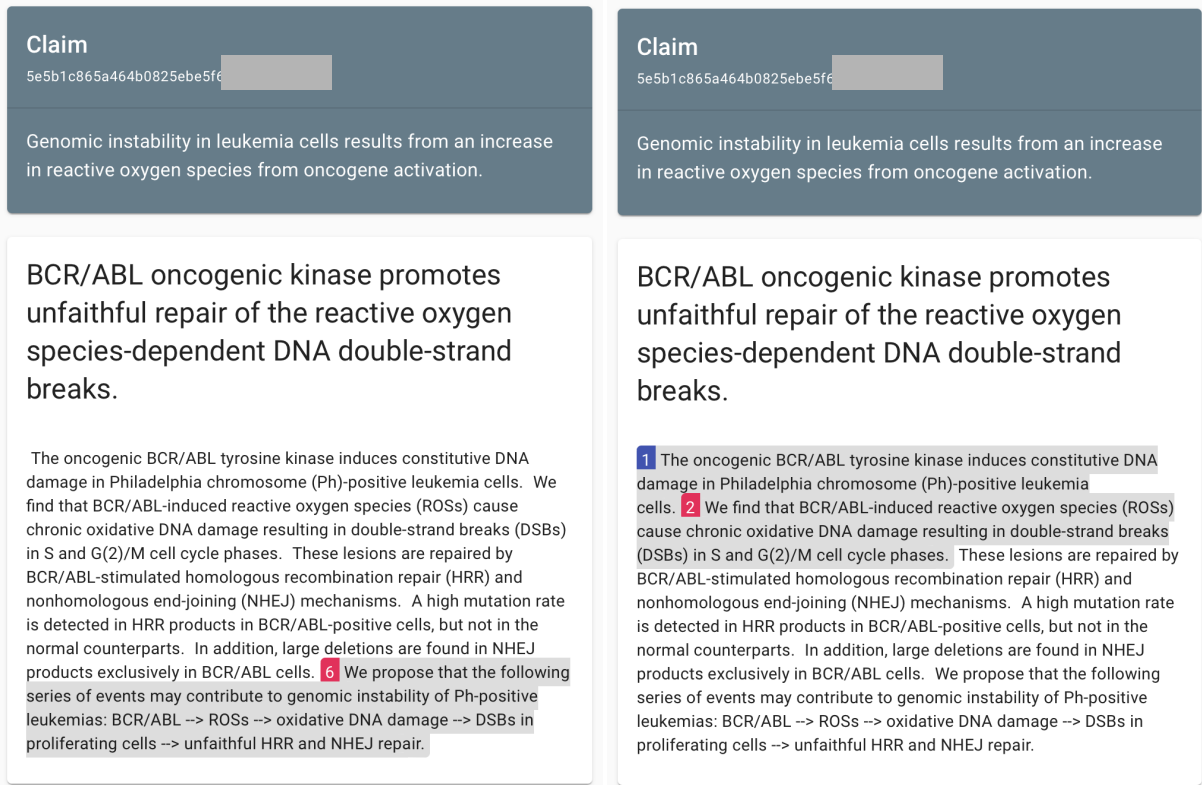
Figure 5 shows a simple example of a claim with two evidence sets. One requires a supplemental sentence, the other does not.

• <b>claim:</b> Aspirin reduces headaches in people genetically predisposed to migraines
• <b>Evidence 1 (primary):</b> In conclusion, aspirin is an effective treatment for headaches in individuals with a genetic predisposition to migraines.
• <b>Evidence 2 (supplemental):</b> We conducted a trial to evaluate the effect of aspirin on headaches in individuals genetically prone to migraines.
• <b>Evidence 2 (primary):</b> The intervention decreased headache frequency by 50% (p = 0.02).

Figure 5: Two example evidence sets. The first evidence set is just a single sentence that is a direct re-phrase of the claim. The second evidence set also supports the claim, but requires collecting sentences from two locations in the document. We want *both* types of evidence sets, not just the obvious re-phrases.

Figure 6 provides a real example of multiple evidence sets, supplemental and primary sentences.

You do not need to provide evidence to support the use of acronyms or synonyms. For instance, if your claim discusses “Human immunodeficiency Virus” it’s fine for an evidence sentence to use “HIV” instead.



(a) First evidence set.

(b) Second evidence set.

Figure 6: Two evidence sets, each providing full support for the claim. The first evidence set has a single primary sentence (see Section 3.1). The second has a primary sentence that states the finding about ROS but doesn't mention leukemia. We add a supplemental sentence to establish that the study is about leukemia.

If it seems like an abstract doesn't provide any support for your claim, make sure that the abstract isn't just using a different synonym or abbreviation. See Figure 7 for an example.

**Claim**

5e5b1c815a464b0825ebe5a [redacted]

Tirasemtiv targets fast-twitch muscle.

### Activation of fast skeletal muscle troponin as a potential therapeutic approach for treating neuromuscular diseases

Limited neural input results in muscle weakness in neuromuscular disease because of a reduction in the density of muscle innervation, the rate of neuromuscular junction activation or the efficiency of synaptic transmission. We developed a small-molecule fast-skeletal-troponin activator, CK-2017357, as a means to increase muscle strength by amplifying the response of muscle when neural input is otherwise diminished secondary to neuromuscular disease. Binding selectively to the fast-skeletal-troponin complex, CK-2017357 slows the rate of calcium release from troponin C and sensitizes muscle to calcium. As a consequence, the force-calcium relationship of muscle fibers shifts leftwards, as does the force-frequency relationship of a nerve-muscle pair, so that CK-2017357 increases the production of muscle force in situ at sub-maximal nerve stimulation rates. Notably, we show that sensitization of the fast-skeletal-troponin complex to calcium improves muscle force and grip strength immediately after administration of single doses of CK-2017357 in a model of the neuromuscular disease myasthenia gravis. Troponin activation may provide a new therapeutic approach to improve physical activity in diseases where neuromuscular function is compromised.

**Submission 1**

Thu Feb 27 at 06:39 PM (25 days ago) [redacted]

---

**Accepted**

Accepted by reviewer

Mon Mar 2 at 01:00 PM (21 days ago)

---

**Evidence Set 1**

Binding selectively to the fast-skeletal-troponin complex, CK-2017357 slows the rate of calcium release from troponin C and sensitizes muscle to calcium.

Support: **Fully Support**

---

**Evidence Set 2**

We developed a small-molecule fast-skeletal-troponin activator, CK-2017357, as a means to increase muscle strength by amplifying the response of muscle when neural input is otherwise diminished secondary to neuromuscular disease.

Support: **Fully Support**

Figure 7: The abstract contains two sentences which each provide full support for the claim. Identifying these sentences as evidence requires knowing that Tirasemtiv is also known as CK-2017357. This information doesn't need to be in the abstract for you to assign full support. I just googled it.

## 3.2 Evidence direction

There are two **directions** that evidence sets can have with respect to a claim.

- **Supports:** You have more reason to believe the claim is **true** after reading the text in the evidence set.
- **Contradicts:** You have more reason to believe the claim is **false** after reading the text in the evidence set.

Keep an eye out for double and triple negatives in evidence text, like **We failed to conclude that aspirin cures headaches.**

## 3.3 Evidence strength

There are two **strengths** that evidence sets can have. Always try to find *all* evidence sets in the abstract. It's fine to have one evidence set that provides full support, and another that provides partial support for the same claim.

We're still in the process of figuring out what full and partial contradiction should look like. There shouldn't be many contradictions in the data yet. If you find one, just mark it as full contradiction and flag it. We'll revise these guidelines later.

To determine full vs. partial *support*, remember these two rules, which are discussed in detail below:

- If the evidence supports a *special case* of the claim (or the exact claim), support is *full*.
- If the evidence supports a *generalization* of the claim (but *not* the exact claim), support is *partial*.

### 3.3.1 Full support

Mark "Full Support" if the sentences in the evidence set provide reasonable support that the claim is true. You should withhold scientific skepticism when making your support decision – for instance, whether or not the methodology in the paper seems flawed shouldn't effect your conclusion. Just take the author's words at face value. See Fig. 1 and Fig. 6 for an example of two evidence sets providing full support.

**Evidence supports a special case of the claim** In Figure 8, the evidence provides support for a *special case* of the claim. Mark these situations as "Full Support". When someone writes "smoking causes cancer", they almost always mean "smoking causes some forms of cancer", not "smoking causes every conceivable form of cancer". For a real example of full support where the evidence discusses a special case of the claim, see Figure 9.

- **Claim:** Smoking causes cancer.
- **Evidence:** Smoking causes lung cancer in men.
- **Support:** Full Support.

Figure 8: This evidence supports a specific case of the claim. This should be marked as *full support*.

**Claim**

5e630276a575092b4d8f2bb3 [redacted]

Energy balance requires hypothalamic glutamate neurotransmission.

**Synaptic glutamate release by ventromedial hypothalamic neurons is part of the neurocircuitry that prevents hypoglycemia.**

The importance of neuropeptides in the hypothalamus has been experimentally established. Due to difficulties in assessing function in vivo, the roles of the fast-acting neurotransmitters glutamate and GABA are largely unknown. Synaptic vesicular transporters (VGLUTs for glutamate and VGAT for GABA) are required for vesicular uptake and, consequently, synaptic release of neurotransmitters. Ventromedial hypothalamic (VMH) neurons are predominantly glutamatergic and express VGLUT2. To evaluate the role of glutamate release from VMH neurons, we generated mice lacking VGLUT2 selectively in SF1 neurons (a major subset of VMH neurons). These mice have hypoglycemia during fasting secondary to impaired fasting-induced increases in the glucose-raising pancreatic hormone glucagon and impaired induction in liver of mRNAs encoding PGC-1alpha and the gluconeogenic enzymes PEPCK and G6Pase. Similarly, these mice have defective counterregulatory responses to insulin-induced hypoglycemia and 2-deoxyglucose (an antimetabolite). **8** Thus, glutamate release from VMH neurons is an important component of the neurocircuitry that functions to prevent hypoglycemia.

**Submission 1**

Sun Mar 15 at 05:42 PM (8 days ago) [redacted] ^

**Review Pending** 🗑️

REVISE NEEDED
REJECT
ACCEPT

Feedback

SUBMIT

**Evidence Set 1**

Thus, glutamate release from VMH neurons is an important component of the neurocircuitry that functions to prevent hypoglycemia.

Support: **Fully Support**

Figure 9: The evidence shows that hypothalamic glutamate release is necessary to prevent hypoglycemia. Since hypoglycemia is a form of energy imbalance, the evidence supports a specific case of the claim. Mark “Full Support”.

### 3.3.2 Partial support

The sentences in the evidence set provide some information, but there are conditions or details stated in the claim that are not satisfied by the evidence set.

- For each partial evidence set, **edit the claim** so that the evidence set provides full support (or contradiction) for the edited claim.
- Make as few changes to the original claim as possible.
- As a check on claim editing, ask yourself “If my re-written claim were True, would that make me more confident in the original claim also being True?” If the answer is “no”, you’ve edited the claim too much. Switch your label to “no evidence”.
- If the claim contains a typo, don’t select “partial support / contradiction” to fix the typo. Instead flag the claim, mention the typo in the note, and find evidence as usual.

The following pages have some examples of situations leading to partial support.

**Evidence supports a generalization of the claim** Figure 10 shows a case where the evidence supports a *generalization* of the claim. This should be marked as partial support, since evidence that smoking causes *some form of cancer* doesn’t imply that smoking causes pancreatic cancer in men *specifically*. Note that this is the opposite of the situation in Figure 8.

- |   |
|---|
| <ul style="list-style-type: none"><li>• <b>Claim:</b> Smoking causes pancreatic cancer in men.</li><li>• <b>Evidence:</b> Smoking causes cancer.</li><li>• <b>Support:</b> Partial support.</li></ul> |
|---|

Figure 10: This evidence supports a generalization of the claim. This should be marked as partial support, and the claim should be rewritten “Smoking causes cancer”.

See Figures 11 and 12 for two real examples of partial support.

**Prefer “full support” over “partial support” when plausible** Give the author the benefit of the doubt. For instance, if it seems plausible that a paper author is using **stem cell** as a shorthand (or *synonym*) for **induced pluripotent stem cell**, it’s fine to use a sentence that discusses stem cells as full support for a claim discussing iPSC’s in particular. Reserve “partial” for cases where this clearly isn’t the case. See 13 for an example of a judgement call like this.



### Claim

5e5b1c8a5a464b0825ebe633 [redacted]

Drosophila supracellular actomyosin structures are found at boundaries in wing imaginal discs.

---

### An actomyosin-based barrier inhibits cell mixing at compartmental boundaries in Drosophila embryos

Partitioning tissues into compartments that do not intermix is essential for the correct morphogenesis of animal embryos and organs. Several hypotheses have been proposed to explain compartmental cell sorting, mainly differential adhesion, but also regulation of the cytoskeleton or of cell proliferation. Nevertheless, the molecular and cellular mechanisms that keep cells apart at boundaries remain unclear. **4** Here we demonstrate, in early *Drosophila melanogaster* embryos, that actomyosin-based barriers stop cells from invading neighbouring compartments. Our analysis shows that cells can transiently invade neighbouring compartments, especially when they divide, but are then pushed back into their compartment of origin. Actomyosin cytoskeletal components are enriched at compartmental boundaries, forming cable-like structures when the epidermis is mitotically active. When MyoII (non-muscle myosin II) function is inhibited, including locally at the cable by chromophore-assisted laser inactivation (CALI), in live embryos, dividing cells are no longer pushed back, leading to compartmental cell mixing. We propose that local regulation of actomyosin contractibility, rather than differential adhesion, is the primary mechanism sorting cells at compartmental boundaries.

Submission 1  
Thu Feb 27 at 07:50 PM (25 days ago) [redacted] ▼

Submission 2  
Wed Mar 4 at 05:10 PM (19 days ago) [redacted] ▲

Review Pending 🗑️

REVISE NEEDED

REJECT

ACCEPT

Feedback

SUBMIT

#### Evidence Set 1

Here we demonstrate, in early *Drosophila melanogaster* **P** embryos, that actomyosin-based barriers stop cells from invading neighbouring compartments.

Support: **Partially Support**

Rewrite: *Drosophila* supracellular actomyosin structures are found at compartment boundaries.

#### Evidence Set 2

Actomyosin cytoskeletal components are enriched at **P** compartmental boundaries, forming cable-like structures when the epidermis is mitotically active.

Support: **Partially Support**

Rewrite: *Drosophila* supracellular actomyosin structures are found at compartment boundaries.

Figure 11: Support is *partial* because the evidence provides support for actomyosin structures at compartment boundaries, but not at the boundaries of wing imaginal discs specifically. Edit the claim to discuss compartment boundaries generally.

### Claim

5e5b1c875a464b0825ebe5fe [redacted]

Loss of myoepithelial cells promote the transition of ductal carcinoma in situ to invasive carcinoma in breast cancer xenograft models.

Submission 1  
Thu Feb 27 at 07:02 PM (25 days ago) [redacted]

Accepted  
Sat Mar 21 at 04:42 PM (2 days ago) [redacted]

### Regulation of in situ to invasive breast carcinoma transition.

**1** The transition of ductal carcinoma in situ (DCIS) to invasive carcinoma is a poorly understood key event in breast tumor progression. Here, we analyzed the role of myoepithelial cells and fibroblasts in the progression of in situ carcinomas using a model of human DCIS and primary breast tumors. Progression to invasion was promoted by fibroblasts and inhibited by normal myoepithelial cells. **4** Molecular profiles of isolated luminal epithelial and myoepithelial cells identified an intricate interaction network involving TGFbeta, Hedgehog, cell adhesion, and p63 required for myoepithelial cell differentiation, the elimination of which resulted in loss of myoepithelial cells and progression to invasion.

#### Evidence Set 1

**S** The transition of ductal carcinoma in situ (DCIS) to invasive carcinoma is a poorly understood key event in breast tumor progression.

**P** Molecular profiles of isolated luminal epithelial and myoepithelial cells identified an intricate interaction network involving TGFbeta, Hedgehog, cell adhesion, and p63 required for myoepithelial cell differentiation, the elimination of which resulted in loss of myoepithelial cells and progression to invasion.

Support: **Partially Support**  
 Rewrite: carcinoma in situ to invasive carcinoma in a breast cancer model system.

#### Evidence Set 2

**S** The transition of ductal carcinoma in situ (DCIS) to invasive carcinoma is a poorly understood key event in breast tumor progression.

**P** Progression to invasion was promoted by fibroblasts and inhibited by normal myoepithelial cells.

Support: **Partially Support**  
 Rewrite: carcinoma in situ to invasive carcinoma in a breast cancer model system.

Figure 12: Support is *partial* because the abstract doesn't ever mention xenograft models. Edit the claim by replacing "xenograft models" with "model system" which is more general.

### Claim

5e5b1c925a464b0825ebe69e · [redacted]

NLRP12 is a negative regulator of T-cell receptor activation.

## The NLRP12 Sensor Negatively Regulates Autoinflammatory Disease by Modulating Interleukin-4 Production in T Cells.

Missense mutations in the nucleotide-binding oligomerization domain (NOD)-like receptor pyrin domain containing family of gene 12 (Nlrp12) are associated with periodic fever syndromes and atopic dermatitis in humans. **2** Here, we have demonstrated a crucial role for NLRP12 in negatively regulating pathogenic T cell responses. Nlrp12(-/-) mice responded to antigen immunization with hyperinflammatory T cell responses. Furthermore, transfer of CD4(+)CD45RB(hi)Nlrp12(-/-) T cells into immunodeficient mice led to more severe colitis and atopic dermatitis. NLRP12 deficiency did not, however, cause exacerbated ascending paralysis during experimental autoimmune encephalomyelitis (EAE); instead, Nlrp12(-/-) mice developed atypical neuroinflammatory symptoms that were characterized by ataxia and loss of balance. Enhanced T-cell-mediated interleukin-4 (IL-4) production promotes the development of atypical EAE disease in Nlrp12(-/-) mice. These results define an unexpected role for NLRP12 as an intrinsic negative regulator of T-cell-mediated immunity and identify altered NF-κB regulation and IL-4 production as key mediators of NLRP12-associated disease.

### Submission 1

Sun Mar 8 at 02:37 PM (15 days ago) [redacted]

## Review Pending

REVISE NEEDED
REJECT
ACCEPT

Feedback

SUBMIT

#### Evidence Set 1

**P** Here, we have demonstrated a crucial role for NLRP12 in negatively regulating pathogenic T cell responses.

Support: **Partially Support**

Rewrite: NLRP12 is a negative regulator of T-cell immune response.

#### Evidence Set 2

**P** These results define an unexpected role for NLRP12 as an intrinsic negative regulator of T-cell-mediated immunity and identify altered NF-κB regulation and IL-4 production as key mediators of NLRP12-associated disease.

Support: **Partially Support**

Rewrite: NLRP12 is a negative regulator of T-cell immune response.

Figure 13: This example is a judgment call. The abstract mentions T-cell response, but not T-cell receptor activation specifically. If it feels plausible that these two terms could be used interchangeably, mark “full”. If you don’t think so, mark “partial”. Two people could reasonably disagree about this. Just use your best judgment.

**Mis-written claims** In Figure 14, partial evidence occurs because the claim states facts about two separate entities. This is a mistake, claims should only discuss a single entity. Do your best to annotate examples like this, but also flag them.

**Claim**

5e5b1c985a464b0825ebe6f6 [redacted]

SIRT1 and DNMT1 are recruited to hydrogen peroxide-induced double strand breaks.

**Oxidative damage targets complexes containing DNA methyltransferases, SIRT1, and polycomb members to promoter CpG Islands.**

Cancer cells simultaneously harbor global losses and gains in DNA methylation. **2** We demonstrate that inducing cellular oxidative stress by hydrogen peroxide treatment recruits DNA methyltransferase 1 (DNMT1) to damaged chromatin. DNMT1 becomes part of a complex(es) containing DNMT3B and members of the polycomb repressive complex 4. Hydrogen peroxide treatment causes relocalization of these proteins from non-GC-rich to GC-rich areas. Key components are similarly enriched at gene promoters in an in vivo colitis model. Although high-expression genes enriched for members of the complex have histone mark and nascent transcription changes, CpG island-containing low-expression genes gain promoter DNA methylation. Thus, oxidative damage induces formation and relocalization of a silencing complex that may explain cancer-specific aberrant DNA methylation and transcriptional silencing.

**Submission 1**

Tue Mar 3 at 08:56 AM (20 days ago) [redacted]

**Review Pending**

REVISE NEEDED

REJECT

ACCEPT

Feedback

**SUBMIT**

**Evidence Set 1**

We demonstrate that inducing cellular oxidative stress by hydrogen peroxide treatment recruits DNA methyltransferase 1 (DNMT1) to damaged chromatin.

Support: **Partially Support**

Rewrite: DNMT1 is recruited to hydrogen peroxide-induced double strand breaks.

Figure 14: The claim in this example discusses two different genes. The abstract has evidence for one of them. Edit the claim to remove the mention of SIRT1, and annotate the example as “partial support”. But also *flag this example*, since this sentence is making a claim about two distinct proteins (instead of just one).

### 3.3.3 Contradiction

Figure 15 shows an example of full contradiction. We'll come up with more explicit rules of thumb concerning partial vs. full and update this document. For now, if you find an example of contradiction, just mark it as full and flag it.

The screenshot displays a clinical trial record interface. On the left, a dark blue box labeled 'Claim' contains the text: 'Metastatic colorectal cancer treated with a single agent fluoropyrimidines resulted in similar efficacy and better quality of life when compared with oxaliplatin-based chemotherapy in elderly patients.' Below this, the title of the trial is shown: 'Chemotherapy options in elderly and frail patients with metastatic colorectal cancer (MRC FOCUS2): an open-label, randomised factorial trial'. The background text includes a 'BACKGROUND' section and a truncated 'METHODS' section. On the right, a green box labeled 'Submission 1' shows the date 'Fri Feb 28 at 10:14 AM (24 days ago)'. Below it, a green box labeled 'Accepted' shows the date 'Mon Mar 2 at 01:12 PM (21 days ago)'. A white box titled 'Evidence Set 1' contains two paragraphs of text. The first paragraph, marked with a red 'P' icon, states: 'On balance, a combination including oxaliplatin was preferable to single-agent fluoropyrimidines, although the primary endpoint of PFS was not met.' The second paragraph, marked with a blue 'S' icon, states: 'METHODS We undertook an open, 2 x 2 factorial trial in 61 UK centres for patients with previously untreated advanced colorectal cancer who were considered unfit for full-dose chemotherapy.' Below the evidence set, the text 'Support: Fully Contradict' is displayed.

Figure 15: The claim asserts that fluoropyriminidines were superior to oxaliplatin, but the evidence document says the opposite. The abstract is truncated, but the evidence set can be seen on the right side of the figure.

### 3.4 Special cases

Two special cases came up during pilot annotations.

#### 3.4.1 Dealing with numbers

When fact-checking numbers, here are some rules of thumb.

1. If the evidence is very close to an exact match for the number in the claim, mark full support.
2. If the evidence is reasonably close but not exact, mark partial support.
3. If the evidence isn't close, mark "contradicts", even if the evidence number isn't a logical contradiction for the claim number. See Figure 16 for an example.

- |  |
|--|
| <ul style="list-style-type: none"><li>• <b>Claim:</b> 20,000 Americans have asthma.</li><li>• <b>Full support:</b> 20,011 Americans have asthma.</li><li>• <b>Partial:</b> 18,102 Americans have asthma.</li><li>• <b>Contradict:</b> 150,000 Americans have asthma.</li></ul> |
|--|

Figure 16: Examples of numerical evidence. For the "Contradict" example, note that if 150,000 Americans have asthma, it is also technically true that 20,000 Americans do. But that's not really the intent of the claim.

### 3.4.2 Counterfactuals and hypotheticals

In Figure 17, the claim says Mice lacking molecule X get a disease, and the evidence says Giving mice molecule X gets rid of the disease. Simple counterfactuals like this can be used as evidence, but make sure to flag them and add a note. Anything more complicated than this example should probably not be used as evidence. If you find yourself constructing an elaborate argument for why a sentence serves as evidence, don't use it.

**Claim**  
5e5b1c8d5a464b0825ebe655

Mice that lack Interferon- $\gamma$  or its receptor are highly susceptible to experimental autoimmune myocarditis.

#### the Development of Autoimmune Myocarditis in Mice by an

**BACKGROUND** Interleukin (IL)-12 exerts a potent proinflammatory effect by stimulating T-helper (Th) 1 responses. This effect is believed to be mediated primarily through the activation of STAT4 and subsequent production of interferon (IFN)- $\gamma$ . **Methods and Results-** We examined the role of IL-12 receptor (IL-12R) signaling in the development of murine experimental autoimmune myocarditis (EAM) induced by cardiac myosin immunization. Both IL-12R $\beta$ 1-deficient mice and STAT4-deficient mice were resistant to the induction of myocarditis. Treatment with exogenous IL-12 exacerbated disease. We questioned whether IFN- $\gamma$  is required for the disease-promoting activity of IL-12. On the contrary, we found that IFN- $\gamma$  suppresses EAM. Lack of IFN- $\gamma$  due to either depletion with an antibody or a genetic deficiency exacerbated myocarditis. Splens from IFN- $\gamma$ -deficient mice immunized with cardiac myosin showed increased cellularity; greater numbers of CD3+, CD4+, CD8+, and IL-2-producing cells; and heightened ability to produce cytokines on stimulation in vitro. **10** Treatment of mice with recombinant IFN- $\gamma$  suppressed the development of myocarditis.

**CONCLUSIONS** IL-12/IL-12R/STAT4 signaling promotes the development of EAM. In contrast, IFN- $\gamma$  plays a protective role. The disease-limiting effects of IFN- $\gamma$  might be explained by its ability to control the expansion of activated T lymphocytes.

**SUBMIT**

#### Evidence Set 1

**S** CONCLUSIONS IL-12/IL-12R/STAT4 signaling promotes the development of EAM.

**P** In contrast, IFN- $\gamma$  plays a protective role.

Support: Fully Support

#### Evidence Set 2

**P** Treatment of mice with recombinant IFN- $\gamma$  suppressed the development of myocarditis.

Support: Fully Support

#### Evidence Set 3

**P** Lack of IFN- $\gamma$  due to either depletion with an antibody or a genetic deficiency exacerbated myocarditis.

Support: Fully Support

#### Evidence Set 4

**P** On the contrary, we found that IFN- $\gamma$  suppresses EAM.

Support: Fully Support

Figure 17: Counterfactual evidence can be used in simple cases.

### 3.5 Difficult examples

Sometimes, an example is just really difficult. When this happens, just do your best, flag the example, and move on. Don't waste too much time on any single claim. Below are some links to really challenging examples, where there was substantial disagreement among three expert annotators.

**The whole abstract is evidence** In Figures 18, a big chunk of the abstract could plausibly be used as evidence. Just do something reasonable, flag it as difficult, and move on.

The image shows a screenshot of a scientific database interface. On the left, there is a dark blue box labeled 'Claim' containing a claim ID '5e5b1c8b5a464b0825ebe63c' and the text 'TNFAIP3 is a glioblastoma tumor enhancer.' Below this is a white box with the title 'Targeting A20 Decreases Glioma Stem Cell Survival and Tumor Growth' and a full abstract text. On the right, there are four evidence sets, each in a white box with a blue border. Each evidence set contains a paragraph of text starting with a red circle containing a white 'P'. Below each evidence set is the text 'Support: Fully Support'.

**Claim**  
5e5b1c8b5a464b0825ebe63c  
TNFAIP3 is a glioblastoma tumor enhancer.

**Targeting A20 Decreases Glioma Stem Cell Survival and Tumor Growth**

Glioblastomas are deadly cancers that display a functional cellular hierarchy maintained by self-renewing glioblastoma stem cells (GSCs). GSCs are regulated by molecular pathways distinct from the bulk tumor that may be useful therapeutic targets. We determined that A20 (TNFAIP3), a regulator of cell survival and the NF-kappaB pathway, is overexpressed in GSCs relative to non-stem glioblastoma cells at both the mRNA and protein levels. To determine the functional significance of A20 in GSCs, we targeted A20 expression with lentiviral-mediated delivery of short hairpin RNA (shRNA). **5** Inhibiting A20 expression decreased GSC growth and survival through mechanisms associated with decreased cell-cycle progression and decreased phosphorylation of p65/RelA. Elevated levels of A20 in GSCs contributed to apoptotic resistance: GSCs were less susceptible to TNFalpha-induced cell death than matched non-stem glioma cells, but A20 knockdown sensitized GSCs to TNFalpha-mediated apoptosis. The decreased survival of GSCs upon A20 knockdown contributed to the reduced ability of these cells to self-renew in primary and secondary neurosphere formation assays. The tumorigenic potential of GSCs was decreased with A20 targeting, resulting in increased survival of mice bearing human glioma xenografts. In silico analysis of a glioma patient genomic database indicates that A20 overexpression and amplification is inversely correlated with survival. Together these data indicate that A20 contributes to glioma maintenance through effects on the glioma stem cell subpopulation. Although inactivating mutations in A20 in lymphoma suggest A20 can act as a tumor suppressor, similar point mutations have not been identified through glioma genomic sequencing: in fact, our data suggest A20 may function as a tumor enhancer in glioma through promotion of GSC survival. A20 anticancer therapies should therefore be viewed with caution as effects will likely differ depending on the tumor type.

**Evidence Set 1**

**P** The decreased survival of GSCs upon A20 knockdown contributed to the reduced ability of these cells to self-renew in primary and secondary neurosphere formation assays.

Support: Fully Support

**Evidence Set 2**

Inhibiting A20 expression decreased GSC growth and survival through mechanisms associated with decreased cell-cycle progression and decreased phosphorylation of p65/RelA.

**P** Elevated levels of A20 in GSCs contributed to apoptotic resistance: GSCs were less susceptible to TNFalpha-induced cell death than matched non-stem glioma cells, but A20 knockdown sensitized GSCs to TNFalpha-mediated apoptosis.

Support: Fully Support

**Evidence Set 3**

The tumorigenic potential of GSCs was decreased with A20 targeting, resulting in increased survival of mice bearing human glioma xenografts.

**P**

Support: Fully Support

**Evidence Set 4**

Although inactivating mutations in A20 in lymphoma suggest A20 can act as a tumor suppressor, similar point mutations

**P** have not been identified through glioma genomic sequencing: in fact, our data suggest A20 may function as a tumor enhancer in glioma through promotion of GSC survival.

Support: Fully Support

Figure 18: I found 6 separate evidence sets for this claim, 4 of which are shown. Note that the first evidence set is counterfactual. It states that glioblastoma stem cells die when A20 is knocked down. Based on this we can infer that expression of A20 promotes glioblastoma cell growth, which supports the claim.



**Confusing claim** In some cases, the claim is just really confusing. Again, do your best but don't let yourself get bogged down with really hard ones.

The screenshot displays a peer review interface. On the left, a dark grey box contains the claim ID '5e5b1c8c5a464b0825ebe64b' and the text: 'Citrullinated proteins externalized in neutrophil extracellular traps act indirectly to perpetuate the inflammatory cycle via induction of autoantibodies.' Below this, a white box contains the title 'NETs are a source of citrullinated autoantigens and stimulate inflammatory responses in rheumatoid arthritis.' and a detailed abstract. The abstract discusses the early events of rheumatoid arthritis (RA) and the role of neutrophil extracellular traps (NETs). It mentions that NETosis is enhanced in RA and that NETs induce inflammatory responses. Two specific points are highlighted with red boxes and numbers: '8 Indeed, during NETosis, neutrophils externalized the citrullinated autoantigens implicated in RA pathogenesis, and anti-citrullinated vimentin antibodies potently induced NET formation.' and '10 In turn, NETs significantly augmented inflammatory responses in RA and OA synovial fibroblasts, including induction of IL-6, IL-8, chemokines, and adhesion molecules.' On the right, a blue header shows 'Submission 1' and 'Sun Mar 8 at 07:45 PM (15 days ago)'. Below this, a blue bar indicates 'Review Pending'. Three buttons are visible: 'REVISE NEEDED' (greyed out), 'REJECT', and 'ACCEPT'. A 'Feedback' text box and a 'SUBMIT' button are also present. A red warning message 'Marked as needs attention!' is displayed. At the bottom, a box titled 'Evidence Set 1' contains two paragraphs of text, each starting with a red 'P' icon, and a 'Support: Fully Support' label.

**Claim**  
5e5b1c8c5a464b0825ebe64b

Citrullinated proteins externalized in neutrophil extracellular traps act indirectly to perpetuate the inflammatory cycle via induction of autoantibodies.

**Submission 1**  
Sun Mar 8 at 07:45 PM (15 days ago)

**Review Pending**

REVISE NEEDED REJECT ACCEPT

Feedback

SUBMIT

**Marked as needs attention!**

**Evidence Set 1**

Indeed, during NETosis, neutrophils externalized the citrullinated autoantigens implicated in RA pathogenesis, and anti-citrullinated vimentin antibodies potently induced NET formation.

In turn, NETs significantly augmented inflammatory responses in RA and OA synovial fibroblasts, including induction of IL-6, IL-8, chemokines, and adhesion molecules.

Support: Fully Support

Figure 19: There's a lot going on in the claim and I had a tough time understanding it.