Detecting Personal Medication Intake in Twitter: Annotation Guidelines

Authors:
Ari Z. Klein, Abeed Sarker, Masoud Rouhizadeh, Karen O'Connor, Graciela Gonzalez

Annotators:
Karen O'Connor, Alexis Upshur

Department of Biostatistics, Epidemiology, and Informatics
Perelman School of Medicine
University of Pennsylvania

Last updated: May 25, 2017

Purpose

The following guidelines will help annotators distinguish between "tweets" that indicate specific instances of the Twitter author's intake of the medication in the tweet (*intake*), tweets that do not specify that the user actually took the medication or when the medication was taken, but are still about the author's intake (*possible intake*), and tweets that do not indicate the author's use of the medication (*no intake*).

Identifying *Intake* Tweets

Annotators should classify tweets as *intake* if they indicate that (i) the medication was actually taken, (ii) the author of the tweet personally took the medication, and (iii) the medication was taken at a specific instance. Some tweets indicate specific instances of personal medication intake in a straightforward way, such as those in which first-person references, temporal markers, and "intake" verb phrases explicitly co-occur in the text of the tweet, as in the following tweets:

(1) I hate fucking Xanax and I just had to take one because I'm literally ab to snap the fuck out

(2) I've been sick for the last 3 days taking Ibuprofen just feel better and to fight Infection "swelling"

(3) It's 7pm here & it's completely throwing me off right now. I needed a Xanex to bring me down to reality.

(4) Sorry for this rant thingy, I took my Vyvanse today lol

Annotators should tag tweets (1), (2), (3), and (4) as *intake* because they explicitly indicate that (i) the medications were taken (*had to take one*, *taking Ibuprofen*, *needed a Xanax, took … Vyvanse*), (ii) it was the authors of the tweets who took the medications (*I*, *I've*, *my*), and (iii) the medications were taken at specific instances of time (*just*, *for the last 3 days*, *7pm*, *right now*, *today*).

Inferring Intake

Most tweets may not as explicitly state that (i) a medication was taken, (ii) the author took it, (iii) and it was taken at a specific time. For example, unlike the tweets above, the following tweet does not contain a verb phrase that explicitly indicates that the medication was taken:

> (5) I've been off adderall about a month now and I'm so much happier, but COMPLETELY useless. I'm like a child again.

Although tweet (5) does not explicitly state that Adderall was *taken*, being *off Adderall* provides enough textual evidence to conclude that it was. Because, in addition to this evidence, tweet (5) explicitly specifies when Adderall was taken—*about a month ago*—and indicates that the author personally took it (*I've*), annotators should tag tweet (5) as *intake*.

We can also infer, from the text of a tweet, that a medication was taken if the tweet indicates (a lack of) an experiential effect of the medication—for example:

> (6) The only thing working right now is ibuprofen but honestly i don't want to have to just rely on that to feel functional
> (7) Tylenol is my bestfriend at the moment
> (8) Thank u Tylenol PM for the wild and crazy dreams last night. Now I have to fig out what's real ?!!??
> (9) Sick and only had a Tylenol PM at work so now i feel better but i am fighting sleep😊

While tweets (6)-(9) do not explicitly state that the medications have been taken, we can infer that they have, from textual evidence that indicates experiential effects of the medications (*working*, *best friend*, *the wild and crazy dreams*, *feel better*, *fighting sleep*). Furthermore, these tweets provide textual evidence that the effects are on the authors (*I*, *my*), and they explicitly specify temporal instances of the effects (*right now*, *at the moment*, *last night*, *now*). Some tweets, however, may indicate (a lack of) an effect of the medication without explicitly indicating that the effect is on the author or when the effect occurred.

Inferring Intake and Time

The following tweets also express (a lack of) an effect of a medication, from which we can reason that the medication was taken, and they explicitly mark that the effect of the medication is on the author. From the combination of this textual evidence, we can reasonably conclude that the authors have taken the medications; however, unlike tweets (6)-(9), the following tweets do not contain surface-level temporal markers of when the medications' effects on the authors occurred:

> (10) the tramadol has Fucked Me Up and Not taken Away My PAIN!!!!!!
> (11) A lot of people hate on prednisone but I feel better already. #stuffworksforme
> (12) Oh my god I want to die never take Cymbalta

Although tweets (10), (11), and (12) do not explicitly specify the instances that the authors took the medications, we will infer from the text that they took the medications recently because the

tweets seem to be about effects that they are *presently* experiencing. If a tweet seems to be indicating a present effect of the medication, and the tweet marks that the effect is on the author, annotators should tag it as *intake*.

Similarly, annotators should tag tweets as *intake* if the authors express that they are *waiting* for an effect of a medication to occur, as in:

(13) Please let this ibuprofen kick in. I'm so sore
(14) this ibuprofen still ain't kicked in my head poundin
(15) Not been to sleep at all. Am in immense discomfort and hoping codeine will kick in so I can sleep for a bit. #SPD

In tweets (13), (14), and (15), the textual evidence that the medications have not *kicked in* implies that the medications were taken and taken recently. In addition, the explicit first-person references (*I'm*, *my*, *I*) indicate that the authors personally took the medications; however, in some cases, the tweets may not explicitly indicate that the medication intake was personally by the author.

Inferring Intake and First-Person Authorship

From the following tweet, we can infer that the medication was taken, and, assuming that it was taken, we know when the medication was taken, but the text does not explicitly indicate that the author of the tweet is the one who took the medication:

(16) In soooo much pain tonight and Tylenol just isn't cutting it. Literally hurting all over

While tweet (16) does not explicitly indicate that the *author* is not experiencing the effect of the Tylenol tonight—a specific time—the high degree of self-presentation in social media allows us to infer that the author is writing about his or her own experience with Tylenol, and, thus, his or her own intake of Tylenol. Annotators should tag tweets as *intake* if the tweets indicate (the lack of) an effect of a medication, explicitly state when the effect is occurring, and do not contain evidence that the effect is being experienced by others (e.g., *you*, *she*, *he*).

Inferring Intake, Time, and First-Person Authorship

The following tweets explicitly indicate neither intake, nor a time of intake, nor that the authors personally took the medications:

(17) Waiting for the ibuprofen to kick in
(18) Prednisone headache! Ahhhh
(19) Ugh. Tylenol as a pregnancy pain reliever is such a joke.

Nonetheless, annotators should still tag tweets (17), (18), and (19) as *intake*. For one, the tweets indicate (not) experiencing an effect of a medication, or waiting to experience an effect, which implies that the medications were taken. In addition, tweet (17) is explicitly occurring in the present (*waiting*), and tweets (18) and (19) also seem to be oriented to the present time, given the

"real-time" nature of social media and the absence of explicit temporal markers of the past or future. From this present orientation, we can infer that the medications were taken recently.

Although the tweets do not explicitly indicate either that the authors are waiting for the medications to *kick in* on themselves, or that they are subjectively experiencing the (lack of) effects of the medications, the first-person nature of social media and the absence of references to others (e.g., *you*, *she*, *he*) give us reason to believe that the authors are writing about their own experiences with the medications, and, thus, their own intake of the medications. While the textual and contextual factors at play in tweets (17), (18), and (19) allow us to infer intake, time, and first-person authorship, we cannot always make this inference, as the guidelines will discuss later.

Inferring Time and First-Person Authorship

While we had to work a little harder to conclude from tweets (17), (18), and (19) that the medications were actually taken, the following tweets provide that information explicitly:

(20) Migraine from hell... Took 6 Motrin and nothing's touching it
(21) *takes Advil just in case*

While tweets (20) and (21) do not explicitly indicate specific times that Motrin and Advil were taken and that the authors personally took the medications, annotators should tag a tweet as *intake* if it contains a verb phrase of medication intake (e.g., *Took 6 Motrin*, *takes Advil*) and does not also contain textual evidence that the medication either was not taken recently or was taken by someone who is not the author of the tweet. In the absence of such temporal or agentive information, the present and personal nature of the social media context allows us to infer that the medication was taken recently, by the author.

Inferring Time

In some cases, a tweet may explicitly indicate that a medication was taken and that it was taken by the author, but it may not specify when it was taken, as in:

(22) I've never been sleepy taking ibuprofen , I took two and I feel like I popped a perc
(23) Popped a Xanax and all I've written so far is: Marketing is cool, I dig that shit. 💀🔫
(24) maybe i'm tired as had 2 tramadol my bk is sore sore sore... #scoliosis

In tweet (22), the personal pronoun, *I*, is explicitly stated in the subject position of the "intake" verb phrase (*I took two*), whereas, in tweets (23) and (24), we can understand *I* to the subject of the "intake" verb phrases (*Popped a Xanax*, *had 2 tramadol*), based on the authors' use of first-person references in other parts of the tweets (*I've*, *I'm*, *my*). While all of the tweets more or less explicitly indicate that the authors have taken the medications, they do not explicitly specify when. We will use the same rationale that we used for tweets (20) and (21) to infer that the authors took the medications recently; that is, tweets (22), (23), and (24) explicitly contain "intake" verb phrases and do not contain textual indications that the medication was not taken recently—for instance, at an unknown time in the past (e.g., *When I would take Xanax…*). Based

on social media's implicit orientation to the present time, annotators should tag tweets (22), (23), and (24) as *intake*.

Inferring First-Person Authorship

In other cases, a tweet may explicitly indicate that a medication was taken and when it was taken, but it may not explicitly state that the author took it, as in:

(25)   Feeling so shitty took Tylenol PM not realizing it was only 6pm
(26)   Just threw back these Xanax

Using the same rationale as earlier, annotators should tag tweets (25) and (26) as *intake* because, even though they do not explicitly indicate that the authors personally took the medications, they do not contain textual evidence that the medications were taken by someone other than the author. Because of the implicit first-person nature of social media, the "intake" verb phrases (*took*, *threw back*), the temporal markers (*6pm*, *just*), and the absence of references to other agents, we can conclude that the authors of tweets personally took the medications at specific instances.

In sum, annotators should tag tweets as *intake* if they indicate a specific time at which the author of the tweet personally used the medication that is mentioned in the tweet.

Identifying *Possible Intake* Tweets

In some cases, it will be rather straightforward that the tweet does not indicate the author's own use of the stated medication (in which case, annotators should tag the tweet as *no intake)*, but there are several other cases that may raise more uncertainty about whether or not the author actually took the medication—for example, tweets that express the intake as a future event:

(27) I want to cry it's that painful 😩gonna take codeine this morning for sure
(28) I guess I will be taking Advil all day haha #FML
(29) About too take this Tylenol PM & pass out
(30) Poorly sick 🤒😷 bed and box set for me tonight. Oh and paracetamol and tea.
        #rocking! 😞😬

In tweets (27)-(30), the authors express that they are going to take the medications in the future (*gonna*, *this morning*, *will be*, *about to*, *tonight*. Annotators should tag tweets that express future intake as *possible intake* because they leave too much room for doubting whether or not the author did end up taking the medications. Tagging tweets (27)-(30) as *possible intake* not only distinguishes them from *intake* cases, in which the textual evidence gives us more confidence that the medications were taken, but also gives us the opportunity the probe the Twitter timelines in which these tweets were found, to see if, for example, subsequent tweets provide more explicit indications that these medications were taken.

The following tweets are similar to tweets that express future medication intake, in that they not contain enough textual evidence that the authors actually took the medications:

(31)   800 mg of Advil cause this headache is real
(32)   Tylenol with codeine for my headache

Annotators should tag tweets such as (31) and (32) as *possible intake*. They should also tag tweets as *poss intake* if they express a present-tense need for a medication, as in:

(33) I need some codeine .
(34) I'm already needing a Tylenol for this vacation and I haven't even gotten there yet
(35) I'm 21 and I'm terrified of the dark. I need a nightlight for fucks sakes. I think I need a Xanax
(36) I need some Tylenol cold & flu
(37) In need of a codeine blunt, and Tony is nowhere near to fulfill my needs.
(38) I need a Xanax like right now

Again, tweets (33)-(38) do not sufficiently indicate that the "needed" medications were actually taken. The distinction between a genuine need for the medication and a rhetorical expression of emotion can be difficult to discern in the absence of additional text or context, so annotators should tag them as *possible intake*, and we can follow-up with the tweets in the users' timelines to see if the "need" was actually fulfilled. Note, however, that the past tense, *needed*, could indicate intake (e.g., *I needed Xanax*, as in tweet (3) earlier).

Annotators should still tag tweets as *poss intake* even if they express a need for *my* [medication], as in:

(39) Nothing pisses me off more than when my cat hides my vyvanse #meow

Even though the use of *my* to modify Vyvanse may more strongly imply that the author does take the medication, alone, it still leaves open too much uncertainty to consider it to be textual evidence of a medication intake. Again, tagging it as *possible intake* would give us the opportunity to probe the timeline to see if we can find a more explicit indication that the author has taken the medication.

Annotators should also tag tweets as *possible intake* if the tweets praise or criticize the medication without explicitly associating it with a concrete effect—for example:

(40)   Codeine is one hell of a medication. ☺☺☺
(41)   baclofen more like bacloFUN
(42)   Why are paracetamol so shit
(43)   found a new love for Klonopin
(44)   I like my Sprite mixed with codeine

As many of earlier tweets illustrate, when tweets describe concrete experiential effects of medications, they tend to provide more evidence that the author took the medications, and at specific times. Tweets (40)-(44) merely praise or criticize the medications, and so provide less textual evidence of intake and time. Such tweets may be pointers to other tweets that more

explicitly indicate personal intake, but, on their own, they only tenuously suggest that the authors may have taken the medications at some point.

Even if tweets contain actions that the author performed that are expressed as out-of-the-ordinary means for obtaining a medication, annotators should tag these tweets as *possible intake*, rather than *intake*. While such expressions, as in the following tweets, can be understood as signaling an immediate need for medication, which might suggest that the author took the medication as soon as he or she obtained it, alone, these expressions, do not provide sufficient textual evidence that the author took the medication:

(45) Up at 3am for an Advil and Popsicles and tea party. My body is a disease hive
(46) When your head is pounding at work and have no choice but to buy $22 Advil from shoppers ☹ rip off !!

While tweets may provide sufficient evidence that the author has personally used the medication, they should be annotated as *possible intake* if they do not also indicate, explicitly or implicitly, a specific instance in which the medication was taken, as in:

(47) 😭😭😭 I never understood why I get so angryyyy omg I was so mellow on Xanax 😵
(48) Up in the middle of the night because everything hurts from prednisone withdrawals. Merry Christmas to my body.

Tweets (47) and (48) provide evidence that the author has taken the medication in the past, but they do not indicate specifically when. Similarly, the following tweets explicitly indicate that the authors frequently use the medications, but they do not specify instances of use:

(49) I pretty much eat Advil like it's candy.💊🙆
(50) I pop Xanax i need more the one imma fein
(51) Idc how sick i get, i just rely on tylenol for everything. I refuse going to the doctor.
(52) If I take any more acetaminophen for my back my liver may explode....

Tweets (49)-(52) do not imply specific instances in which the authors took the medications (in the way that *past-tense* "intake" verb phrases do, for example); rather, they only indicate that the authors take the medications frequently in general. Similarly, some medications *imply* that the authors take the medications frequently, such as *Albuterol* (which is used commonly to treat asthma) in the following tweet:

(53) Perk of having a NP husband: we're at Bartells getting an Albuterol inhaler. (And yes, I already on a 2x daily steroid AND I took Serovent.)

Annotators should tag tweets as *possible intake* that merely indicate frequent use, without specifying an instance in which the author took the medication. Tweet (53), which includes three medications, highlights that it is also important for annotators to pay careful attention to the specified medication of interest for each tweet.

In sum, annotators should tag tweets as *possible intake* if they are about the author's personal intake of the medication, but do not contain sufficient textual evidence that the author actually took the medication or of specifically when the author took it.

Identifying *No Intake* Tweets

In general, annotators should tag tweets as *no intake* if they are not explicitly about the author's own intake of the medication—for example, tweets in which author the suggests that others should take the medication:

(54) @[Username redacted] Mine hurt for days last year!! Take some paracetamol hun ☺

In tweet (54), the suggestion to take the medication might be interpreted as implying that the author has taken the medication; furthermore, the temporal marker, *last year*, not only provides a specific instance in which the author may have taken the medication, but it seems to provide additional evidence that the author has taken the medication. Nonetheless, annotators should tag tweets such as (54), which merely suggest taking a medication, as *no intake* because they do not contain sufficient textual evidence upon which to draw the conclusion that the author has personally taken the medication.

Similarly, annotators should tag tweets as *no intake* if they are about the author's *giving* the medication to someone else, as in:

(55) Gave James 2 ibuprofen pm and I'm being repaid by the sound of him snoring penetrating through my earplugs

While tweets (54) and (55) are clearly not about the authors' own intake of the medications, other cases may not be as obvious—for example, when the tweet contains merely the name of a medication:

(56) @[Username redacted] @[Username redacted] @[Username redacted] @[Username redacted] methadone !

Although tweet (56) might be interpreted as indicating the author's use of Methadone—possibly even the author's *present* use of Methadone—the textual evidence does not seem to support this interpretation more heavily than other possible interpretations. Because tweet (56) does not give us sufficient reason to investigate the author's timeline, annotators should tag it as *no intake*.

An even less obvious case of *no intake* is when the author repeats the speech of someone else, as in:

(57)   @[Username redacted] "I don't either cause these Tylenol aren't doing crap!" Lol

In tweet (57), the quotation marks indicate that these words are probably not the author's; rather, it appears that the author is commenting on these words (*Lol*) and directing this comment to the Twitter user, *@[Username redacted]*, perhaps whose words they are. Therefore, while the tweet

does have a feature of *intake* tweets (i.e., it describes a personal effect of a medication), the quotation marks remove our reason to believe that the author has used the medication. Because of the dialogical nature of social media, annotators should be careful about cases of reported speech. Annotators should tag all cases of reported speech as *no intake*.

In addition to quotation marks, as in tweet (57), using a hyphen at the end of a text, followed by someone's name, is a way of marking reported speech and attributing authorship to others—for example:

(58) I just wanna give a shoutout to adderall for helping me get through the semester - Florida State

Tweet (58) uses a hyphen at the end, followed by *Florida State*, to indicate that this is someone else's tweet and attribute authorship to that author. While the tweet contains personal references (*I*, *me*), temporal markers (*the semester*), and an effect that the medication, *Adderall*, had on the author of origin (*helping me get through the semester*), for the same reason as tweet (58), it should be annotated as *no intake*.

Perhaps the most difficult case of reported speech is when a tweet makes a cultural reference—for example, to song lyrics or a line from a movie—that is about taking a medication, as in:

(59) Sunglasses and Advil 😎
(60) Codeine crazy 🙂
(61) I take Quaaludes 10-15 times a day for my "back pain", Adderall to stay focused, Xanax to take the edge off, pot to mellow me out, cocaine to wake me back up again, and morphine... Well, because it's awesome.
(62) that xanax make you trip, i still ain't got no whip, i still ain't got no friends, at least i got my family.. bullshit.
(63) codeine, xanax keep me tame but nothings like my mary jane
(64) zoloft & paxil & buspar & xanax depakote klonopin ambien prozac atavan calms me when i see the bills these are a few of my favorite pills
(65) Codeine cups paint a picture so vivid

Annotators should tag all cultural references to medications as *no intake*. In addition to drawing upon their own knowledge in order to identify cultural references, annotators should pay attention to (variants of) tweets that seem to occur with a relatively high frequency, and investigate whether they are cultural references; a higher frequency could be a clue that the tweets are not providing evidence of personal medication intake.

To recap, annotators should tag tweets as:
- *intake* if they indicate a specific time at which the author of the tweet personally used the medication;
- *possible intake* if they are about the author's personal intake of the medication, but it is unclear that the author actually took the medication or when the author took it;
- *no intake* if they are not explicitly about the author's personal intake of the medication.