

Detecting Birth Defects on Twitter: Annotation Guidelines

Authors:

Ari Z. Klein, Abeed Sarker, Haitao Cai, Davy Weissenbacher, Graciela Gonzalez

Annotators:

Karen O'Connor, Alexis Upshur

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

Last updated: April 11, 2018

Purpose

The following guidelines will help annotators distinguish three classes of tweets: (1) tweets that indicate that the Twitter user's child has a birth defect ("defect"); (2) tweets that indicate a parent-child relationship but leave doubt about whether the child has a birth defect, or vice versa ("possible defect"); and (3) tweets that provide no reason to believe that the user has a child with a birth defect ("non-defect").

Identifying "Defect" Tweets

Annotators should classify tweets as "defect" if they indicate that (i) an individual who is referenced in the tweet has a birth defect and (ii) the user is the parent of that individual. Some tweets indicate (i) and (ii) explicitly, such as the following tweets:

- (1) **My daughter** [name] **born** on 12/20/15 **with** trisomy 21
- (2) #Microcephaly **was the only symptom at birth** and **our child** went undiagnosed and untreated. Please docs think #CMV!

Tweet (1) explicitly states that (i) an individual was born with trisomy 21 and (ii) she is the child of the user, and (2) explicitly states that (i) an individual has microcephaly and (ii) the user is the parent of that individual.

Other tweets may not as explicitly state that (i) an individual has a birth defect or (ii) the individual is the child of the user, as in:

- (3) **My little miracle**, we are so proud and blessed to have you
#hypoplasticleftheartsyndrome #hlhs #miracle
- (4) Smiling DURING her blood draw. This kid is a pro! #heartdisease #heartmom #hlhs

By using a metaphor, *my...miracle*, to refer to the user's child, (3) does not as explicitly indicate a parent-child relationship as (1) and (2). It also requires piecing together the elements of the tweet to conclude that the child has hypoplastic left heart syndrome.

Requiring even more work, (4) does not modify *kid* with a possessive determiner (e.g., *my*, *our*) to indicate a parent-child relationship, but rather does so by burying this parent-child relationship in the hashtag, *#heartmom*, alongside the other hashtags; nonetheless, (3) and (4) do indicate that (i) an individual has a birth defect and (ii) the user is the parent of that individual, so annotators should classify them as “defect.”

Identifying “Possible Defect” Tweets

In some cases, the tweet may indicate a parent-child relationship but leave open whether the child has a birth defect—for example:

- (5) @[username] help spread awareness for Down Syndrome. We are walking in just two weeks. This is **my sons** donation page.

Although (5) explicitly establishes that the user is the parent of the child that is mentioned in the tweet (*my son*), it is unclear whether the child *has* Down syndrome; based on the information in the tweet, we can only conclude that he is walking to raise money for Down syndrome, not that he is necessarily affected by it. If a tweet indicates a parent-child relationship but does not contain explicit evidence that the child is affected by the birth defect, it should be annotated as “possible defect.”

Annotators should also classify a tweet as “possible defect” if the tweet expresses uncertainty about whether or not the user's child has a birth defect, as in:

- (6) **My baby** could possibly have Down syndrome.. 😊
 (7) Prayers for **my baby** [name] would be appreciated. Doctors are keeping an eye out for hydrocephalus ..

Tweets (6) and (7) indicate that the user is the parent of the individual (*my baby*), but leave doubt about whether the child in (6) has Down syndrome (*could possibly have*) or the child in (7) has hydrocephalus (*keeping an eye out for*).

Alternatively, some tweets may explicitly state that an individual has a birth defect, but leave doubt about whether the individual is a child of the user—for example:

- (8) We knew early on in pregnancy that [name] **had** a series of congenital heart defects, known collectively as Tetralogy of Fallot.
 (9) **Baby has** a crystal skull shaped elongated skull.
 (10) What it's like to find out your **baby has** a cleft lip #LotsToLearn #baby

While (8) contains explicit evidence that the individual has a birth defect, the referent of the individual’s name is ambiguous; we cannot necessarily conclude that the individual is the child of the user. The user may or may not be the parent of individual, so annotators should classify (8) as “possible defect.” We may find evidence in the user’s timeline that the individual is (not) the name of the user’s child.

Because (9) omits a determiner—possessive (e.g., *my*, *her*, *their*) or demonstrative (e.g., *this*)—before *baby*, it is unclear whether the user is referring to her own child or someone else’s. On the other hand, (10) does contain a possessive determiner before *baby*, but it is unclear whether *your* is being used in this case to imply a self-reference—that is, to indicate that the *user’s* child has a cleft lip. (If *your* is an obvious reference to others, the tweet should be annotated as “non-defect.”)

Similarly, the following tweet implies (rather than explicitly states) that an individual has a birth defect, but leaves doubt about whether the user is the parent of the individual:

(11) Thought I'd never have to encounter Prune Belly syndrome. Today thought to show me otherwise

In (11), we can infer from *to encounter Prune Belly syndrome* that an individual has a birth defect, but the tweet does not make it clear whether the individual to which the user is referring is the *user’s* child; the user may be the child’s parent, but, alternatively, the user may be a clinician, for example. Because (11) indicates that an individual has a birth defect, but it leaves doubt about the nature of the relationship between the user and the individual, it should be annotated as “possible defect.”

Importantly, while (8), (9), (10), and (11), in different ways, all leave doubt about whether the individual who is referenced in the tweet is the child of the user, they nonetheless provide reason to believe that such a relationship is plausible; this reason seems to be conveyed in the form of an explicit or implicit reference to an individual with whom the user has had a personal experience. We can infer this personal experience from descriptive information in the tweet, which warrants further investigation into whether this is a parent-child relationship.

Further probing into the users’ timelines might provide evidence of whether or not the users are the parents of the individuals in (8), (9), (10), and (11), and whether or not the children in (5), (6), and (7) have a birth defect. To recap, annotators should classify tweets as “possible defect” if they leave doubt about whether an individual is the child of the user or whether that child has a birth defect.

Identifying “Non-Defect” Tweets

Annotators should classify tweets as “non-defect” if the tweets provide no reason to believe that the user is a parent; for example, tweets that do not even refer to specific individuals, as in the following tweets, provide no reason to believe that the user is a parent:

- (12)Gastroschisis looks so scary. Probably shouldn't look at pics on google, but I can't resist.
#hopingforthebest
- (13)@[username] please RT to raise awareness of #CHD & #20weekscan

Because (12) and (13) do not refer to specific individuals, explicitly or implicitly, they fail to even raise doubt about whether the user is a parent; they provide no evidence of a parent-child relationship. Without such evidence, the tweets do not point us to a child who (even possibly) has gastroschisis or a congenital heart defect (CHD); therefore, (12) and (13) should be annotated as “non-defect.”

In other cases, tweets may refer to individuals and even state that the individuals have a birth defect, but still provide no reason to believe that the user is a parent—for example:

- (14)**Kids with** Down Syndrome.
- (15)Share if you love **someone with** Down syndrome!

Tweets (14) and (15) refer to individuals who have Down syndrome, but, because the tweets do not imply that the user has had a personal experience with the referenced individual, they do not provide a reason to believe that the user might be a parent of the individual; therefore, annotators should classify them as “non-defect.”

If a tweet refers to a specific individual whom we can assume is not the child of the user, then, of course, it also provides no reason to believe that the user is a parent of the individual, as in:

- (16)It pisses me off when people talk about my nephews lip . Like he didn't ask to be **born with** a cleft lip. He's perfect to me & my family 🙌❤️

While (16) refers to a specific individual and states that the individual has a cleft lip, it explicitly encodes that the individual is not the child of the user, but rather the son of the user’s sibling (*my nephew*).

Some tweets may appear to contain evidence that the user is a parent of a referenced individual, but require a little more attention to realize that they actually provide no such evidence, as in:

- (17)South African mom: abortion would've spared **my son suffering from** Down syndrome

Although (17) explicitly indicates a parent-child relationship (*my son*) and links the user’s child to the birth defect (*suffering from*), it also contains evidence that the tweet is not about the *user’s* child, but rather the child of a reported individual (*South African mom*). In this case, the explicit source attribution distinguishes this tweet from “defect” tweets.

Similarly, the following tweet illustrates the importance of examining the context surrounding a birth defect in deciding if the tweet is really indicating that the user is a parent:

(18) Its fun with feet in #Kenya, as **our #clubfoot kids** run and play clubfoot free today!
#runfree2030 #WorldClubfootDay

Although (18) contains evidence that the user is the parent of kids who have a birth defect (*our #clubfoot kids*), the context of the tweet suggests that the referent of *our* is not parents, but rather a sort of charity organization. Because (18) actually provides no reason to believe that the user is a parent, it should be annotated as “non-defect.”

Annotators should also classify tweets as “non-defect” if the tweet explicitly states that the user’s child does not have a birth defect, as in:

(19) **My baby's** Trisomy 18 results came back NEGATIVE! I can't stop crying I'm so happy!

If a tweet merely points out that the user is having a test for a birth defect, but does not indicate that the child might be born with the birth defect, the tweet should be annotated as “non-defect”—for example:

(20) Doctors appt in the morning, hearing **my baby** heartbeat again ;) & down syndrome test...

In some cases, the birth defects themselves may also lead to difficulties in distinguishing a “non-defect” tweet from the other classes of tweets—for example:

(21) I hate that I have to deal with an umbilical hernia the entire time that I'm pregnant. It scares me so much. 😞

An umbilical hernia can occur in mothers as well as in babies, so, in such cases, annotators will have to decide whether the parent or the child is the one who is affected. Annotators should use the guidelines for identifying “defect” and “possible defect” tweets to determine if the *child* was (or will be) born with a birth defect. The guidelines would suggest annotating (23) as “non-defect.”