

Tobi

Ungenauigkeiten im Algorithmus

Tobi ist 37 und ein autodidaktisch ausgebildeter Musiker und Technik-Enthusiast. Als Kind wurde bei ihm Diabetes Typ 1 diagnostiziert, als Untersuchungen zeigten, dass er kein Insulin zur Regulierung seines Blutzuckerspiegels produzieren kann. Er ist als Tontechniker tätig und führt ein arbeitsreiches Leben, reist viel herum und arbeitet Nachtschichten. Er genießt seinen spannenden und vielseitigen Job, aber manchmal vergisst er, wann und wieviel Insulin er sich verabreichen muss. Deshalb ist er begeistert, als er eine App für seine Smartwatch entdeckt, die ihn an die Insulineinnahme erinnert. Die App misst sogar seinen Blutzuckerspiegel mit einer neuen sensorbasierten Technologie durch den Schweiß unter der Uhr. Auf Basis dieser Messungen liefert ihm die App personalisierte Empfehlungen bezüglich der richtigen Dosierung. Tobi ist erleichtert, denn er glaubt, dass er mit der App seine Gesundheit nun voll im Griff hat, und seine musikalischen Abenteuer uneingeschränkt genießen kann.

Tobi ist sich allerdings nicht bewusst, dass die Praktik der Glukosemessung über den Schweiß nicht hinreichend getestet wurde und für den medizinischen Gebrauch nicht sicher ist. Die Smartwatch liest seine Blutzuckerwerte falsch aus und die App liefert Tobi falsche Empfehlungen zur Insulindosierung. Eines Tages nimmt Tobi deshalb zu viel Insulin und erleidet einen diabetischen Schock.

Als Tobi sich erholt, schreiben er und seine Ärztin eine Beschwerde über die App an eine Gesundheitsaufsichtsbehörde. Die Behörde beginnt eine Untersuchung sowohl der Smartwatch als auch der App, da es mehrere Beschwerden gab. Die Untersuchung zeigt, dass die Smartwatch aufgrund fehlender systematischer Testungen mit einem adäquaten Forschungsdesign keine korrekten und zuverlässigen Blutzuckermessungen liefern kann. Darüber hinaus zeigt sich, dass dem Entwicklungsteam der App keine medizinischen Fachleute angehören. Die Ergebnisse weisen auch auf gewisse Einseitigkeiten hin, die die von der App generierten Gesundheitsempfehlungen ungenau werden lassen. Die Untersuchung zeigt, dass für die Entwicklung gar nicht zwischen Typ 1 und Typ 2 Diabetes unterschieden wurde. Der von der App verwendete Algorithmus stammt lediglich von Trainingsdaten aus klinischen Studien mit Personen mit Diabetes Typ 2, was dadurch verursacht wird, dass der Körper Glukose nicht verstoffwechseln kann.

Durch die Untersuchung wird auch deutlich, dass das Entwicklungsteam bei der Konzeption der App davon überzeugt war, die Technologie würde bessere Ergebnisse liefern, wenn sie individuelle Daten mit den Daten anderer Nutzenden kombiniert, um eine durchschnittlich gute Insulindosis zu ermitteln. Aus diesem Grund waren die vom Algorithmus generierte Empfehlungen nicht auf Tobis individuellen Bedürfnisse zugeschnitten.

Diskussion: Ungenauigkeiten im Algorithmus

Der Fall von Tobi wirft viele ethische Fragen zur Wirksamkeit, Verlässlichkeit und Sicherheit von mHealth-Technologien auf. In diesem Sinne veranschaulicht er die große Bedeutung einer adäquaten Entwicklung und Prüfung von Geräten, die eine personalisierte Gesundheitsversorgung und medizinische Beratung versprechen.

Man kann sich nun fragen, welche Kriterien für eine gute Gestaltung von mHealth-Produkten wichtig sind. Tobis Fall verdeutlicht, dass eine gut gestaltete Technologie auf adäquaten und medizinisch verlässlichen Daten entsprechend der Erkrankung beruhen müssen, da sie sonst die Anfälligkeit der Nutzenden verschlimmern und ihnen sogar Schaden zufügen können. Eine gute mHealth-Technologie muss auch gesundheitliche Vorteile bieten, was nur auf Grundlage medizinisch begründeter Erkenntnisse möglich ist. Jemand, der mHealth nutzt, erwartet zu Recht eine korrekte und personalisierte Gesundheitsberatung. Damit eine Technologie diese Erwartungen erfüllt, muss sie Trainings- und Nutzungsdaten auf eine für die jeweilig nutzende Person relevante Weise verarbeiten

können. Dabei sind Kriterien wie der Diabetes-Typ ebenso wichtig wie andere Faktoren, etwa die ethnische Zugehörigkeit oder das Geschlecht. Es ist beispielsweise bekannt, dass in der Vergangenheit viele klinische Studien mit hellhäutigen Männern durchgeführt wurden. Dadurch sind klinische Studien verzerrt oder „gebiased“, denn die Ergebnisse sind zunächst nur für hellhäutige Männer relevant.

Es ist schon lange bekannt, dass das in der medizinischen Praxis zu verzerrten Annahmen und Gesundheitsempfehlungen z.B. für Frauen oder Personen mit dunklerer Hautfarbe führen kann. Wenn solche verzerrten Ergebnisse dann Algorithmen für mHealth-Technologien zugrunde gelegt werden, kann sich auch in der digitalisierten Gesundheit dieser Bias weiter fortsetzen und verfestigen. In der Praxis kann es zu Fehlangaben der mHealth-Technologie für Personen führen, die z.B. nicht weiß und männlich sind.

Daher ist es wichtig, dass klinische Leitlinien und mHealth-Technologien sensibel sind für einen möglichen algorithmischen Bias z.B. aufgrund von Hautfarbe oder Geschlecht. Es sollte Mechanismen geben, wie solche Verzerrungen gefunden und reduziert werden können. Auf diese Weise kann die Therapie verbessert und das Fortschreiten der Krankheit und die Entwicklung von Komplikationen reduziert werden.

Um eine korrekte Diagnose und eine gute personalisierte Gesundheitsversorgung für alle Menschen zu gewährleisten, sollten mHealth-Technologien auf Daten von verschiedenen Bevölkerungsgruppen basieren. Bei den von diesen Technologien verwendeten Algorithmen ist es absolut entscheidend, jegliche Einseitigkeiten oder Verzerrungen zu eliminieren, die zu falschen oder irreführenden Gesundheitsempfehlungen und möglicherweise auch zu Schäden für die Nutzenden führen können. Wie wir im Fall von Tobi sehen, kann eine falsche Blutzuckermessung in Kombination mit Trainingsdaten, die für einen anderen Diabetestyp relevant sind, dazu führen, dass die von mHealth-Technologien angebotene Gesundheitsberatung zu ungünstigen individuellen Gesundheitsergebnissen führt.

Um Schaden von den Nutzern abzuwenden, müssen mHealth-Technologien adäquat auf die Sicherheit und Wirksamkeit sowie auf die Angemessenheit, Ausgewogenheit und Verlässlichkeit ihrer Algorithmen getestet werden.

LITERATUR

Campolo,

A., Sanfilippo, M., Whittaker, M. and Crawford, C. (2017). AI Now 2017 Report. *AI Now Institute*.

Gale,

E. A. M.; Gillespie, K. M. (2001). Diabetes and gender. *Diabetologia*, 44(1), 3-15.

Sharp,

M., O'Sullivan, D. 2017. Mobile Medical Apps and mHealth Devices: A Framework to Build Medical Apps and mHealth Devices in an Ethical Manner to Promote Safer Use - A Literature Review. *Stud Health Technol Inform*, 135: 363-367.

Schneider,

A. L. C., Lazo M., Ndumele, C. E., Pankow, J. S., Coresh, J., Clark, J. M., Selvin, E. (2013). Liver enzymes, race, gender and diabetes risk: the Atherosclerosis Risk in Communities (ARIC) Study. *Diabetic Medicine*. 30(8): 926-933.

Siddiqui,

M. A., Khan, M. F., Carline, T. E. (2013). Gender Differences in Living with Diabetes Mellitus. *Mater Sociomed*,

25(2),140-142.

Legato,
M.J., Gelzer A., Goland R., Ebner S.A., Rajan, S., Villagra, V.,
Kosowski, V., The Writing Group for The Partnership for
Gender-Specific Medicine. Gender-specific care of the patient with
diabetes: Review and recommendations. (2016). *Gender Medicine*, 3(2),
131-158.

Juutilainen,
A., Kortelainen, S., Lehto, S., Rönnemaa, T., Pyörälä, K,
Laakso, M. (2004). Gender Difference in the Impact of Type 2 Diabetes
on Coronary Heart Disease Risk. *Diabetes Care* 2004 Dec; 27(12):
2898-2904.

QUELLE

This case is inspired by a New York Times blog. The story raised issues with untested and unreliable apps, including those which are failing to count blood glucose correctly.

Krisch, J. A. (2015). Questioning the Value of Health Apps. *The New York Times*, 16. 3. 2015.