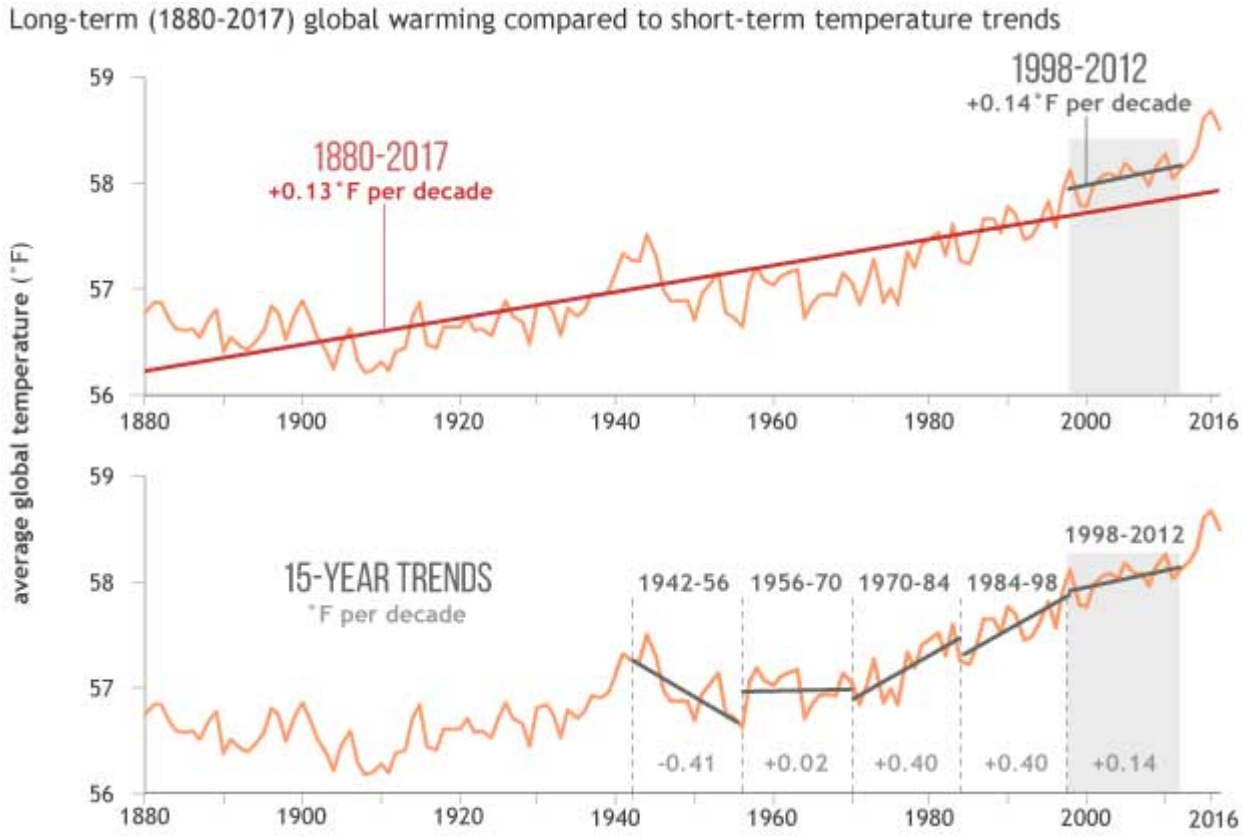
Warum man in Graphiken keine Trendlinien legen sollte

geschrieben von Chris Frey | 9. August 2019

Wir sehen hier bei WUWT (und EIKE Anmerkung des Übersetzers) sehr viele Graphiken – alle Arten von Graphiken von vielen verschiedenen Datensätzen. Beispielhaft sei hier eine allgemein gezeigte Graphik der NOAA gezeigt, die aus einem Stück bei *Climate.gov* stammt und die Bezeichnung trägt „Did global warming stop in 1998?" von Rebecca Lindsey, veröffentlicht am 4. September 2018:



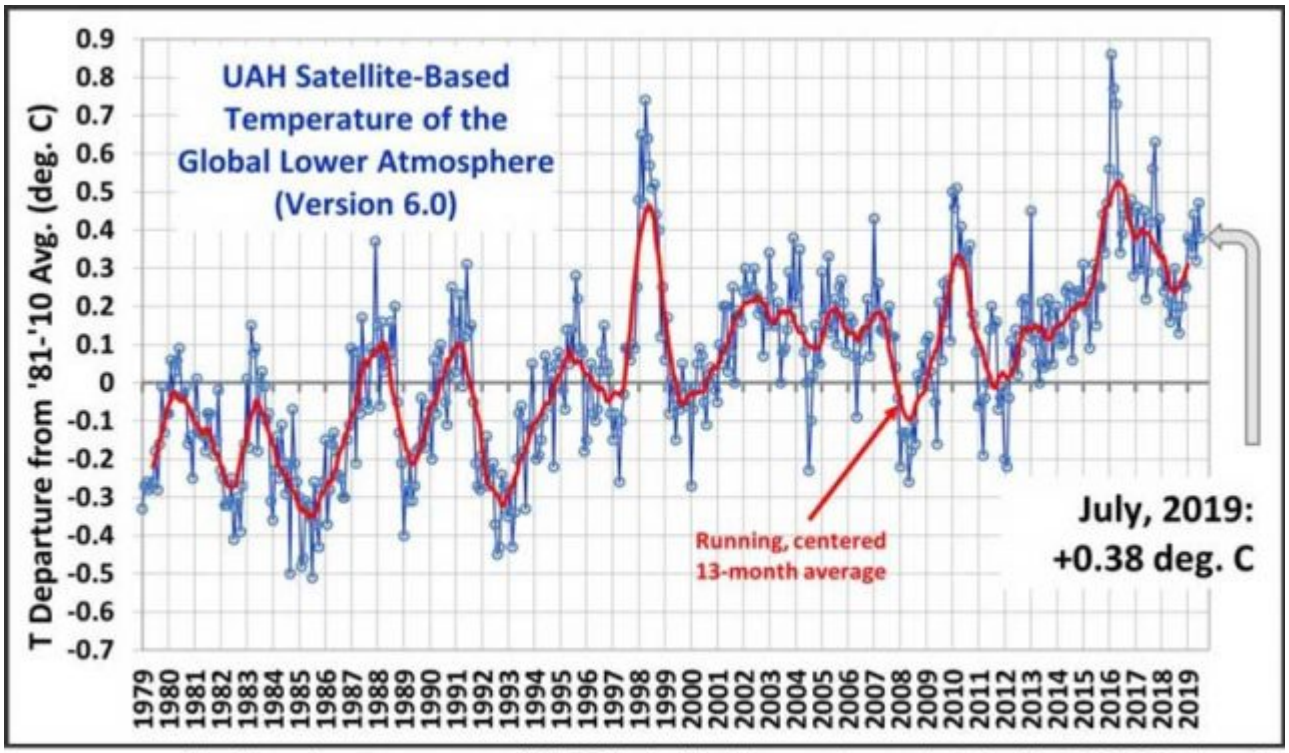
Die Details in dieser Graphik interessieren mich nicht – das Ganze qualifiziert sich selbst als „dümmlich". Die vertikale Skala ist in Grad Fahrenheit angegeben und die gesamte Bandbreite über 140 Jahre liegt größenordnungsmäßig bei 2,5°F oder etwa 1,5°C. Interessant in der Graphik ist die Mühe, „Trendlinien" über die Daten zu legen, um den Leser etwas über die Daten zu vermitteln, was der Autor der graphischen Repräsentation übermitteln will. Dieses „etwas" ist eine Meinung – es ist immer eine Meinung – es ist nicht Teil der Daten.

Die Daten sind die Daten. Wenn man die Daten in eine graphische Darstellung zwängt, hat man bereits Meinung und persönliche Beurteilung einfließen lassen in Gestalt der Wahl von Anfangs- und Endzeitpunkt, vertikalen und horizontalen Skalen und, in diesem Falle, die

Schattierung eines 15-Jahre-Zeitraumes an einem Ende. Manchmal übernimmt irgendeine Software die Entscheidung der vertikalen und horizontalen Skala – und nicht rationale Menschen – was sogar zu noch größerer Konfusion führt und manchmal zu großen Fehlinterpretationen.

Jeder, der die Daten in der oberen Graphik nicht eindeutig erkennt *ohne die Hilfe der roten Trendlinie* sollte sich ein anderes Studienfeld suchen (oder seinem Optiker einen Besuch abstatten). Die untere Graphik ist in ein Propaganda-Werkzeug verwandelt worden mittels Addition von fünf Meinungen in Gestalt von Mini-Trendlinien.

Trendlinien ändern die Daten nicht – sie können lediglich die Sichtweise auf die Daten verändern (hier). Trends können manchmal nützlich sein (bitte mit einem großen *vielleicht* versehen), aber sie bewirken in den Graphiken der NOAA oben nichts Anderes als zu versuchen, den vom IPCC sanktionierten Gedanken des „Stillstands“ zu verunglimpfen. Damit wollen der Autor und die Herausgeber ihre gewünschte Meinung bei *Climate.gov* unterstreichen. Um Rebecca Lindsey aber etwas gerecht zu werden – sie schreibt „wie *viel* langsamer der Anstieg erfolgt, hängt vom Kleingedruckten ab: nämlich welchen Datensatz der globalen Temperatur man betrachtet“ (hier). Dazu hat sie sicher das Recht. Hier folgt die globale mittlere Temperatur der unteren Troposphäre von Spencer an der UAH:



http://www.drroyspencer.com/2019/08/uah-global-temperature-update-for-july-2019-0-38-deg-c/

Man braucht hier keinerlei Trendlinien, um den Stillstand zu erkennen, welcher sich vom Ende des Super El Nino 1998 bis zum Beginn des El Nino 2015-2016 erstreckt. Dies illustriert zweierlei: Hinzugefügte Trendlinien liefern zusätzliche Informationen, die nicht Bestandteil des Datensatzes sind, und es ist wirklich wichtig zu wissen, dass es für

jedwedes wissenschaftliches Konzept mehr als nur einen Datensatz gibt – mehr als eine Messung – und es ist entscheidend wichtig zu wissen „*What Are they Really Counting?*“, wobei der zentrale Punkt Folgender ist:

Also haben wir bei allen gezeigten Messungen, die uns als Informationen angeboten werden, besonders wenn sie von einer behaupteten Signifikanz begleitet werden – wenn man uns also sagt, dass diese Messungen/Zahlen dies oder das bedeuten – eine grundlegende Frage: ***Was genau registrieren sie da eigentlich?***

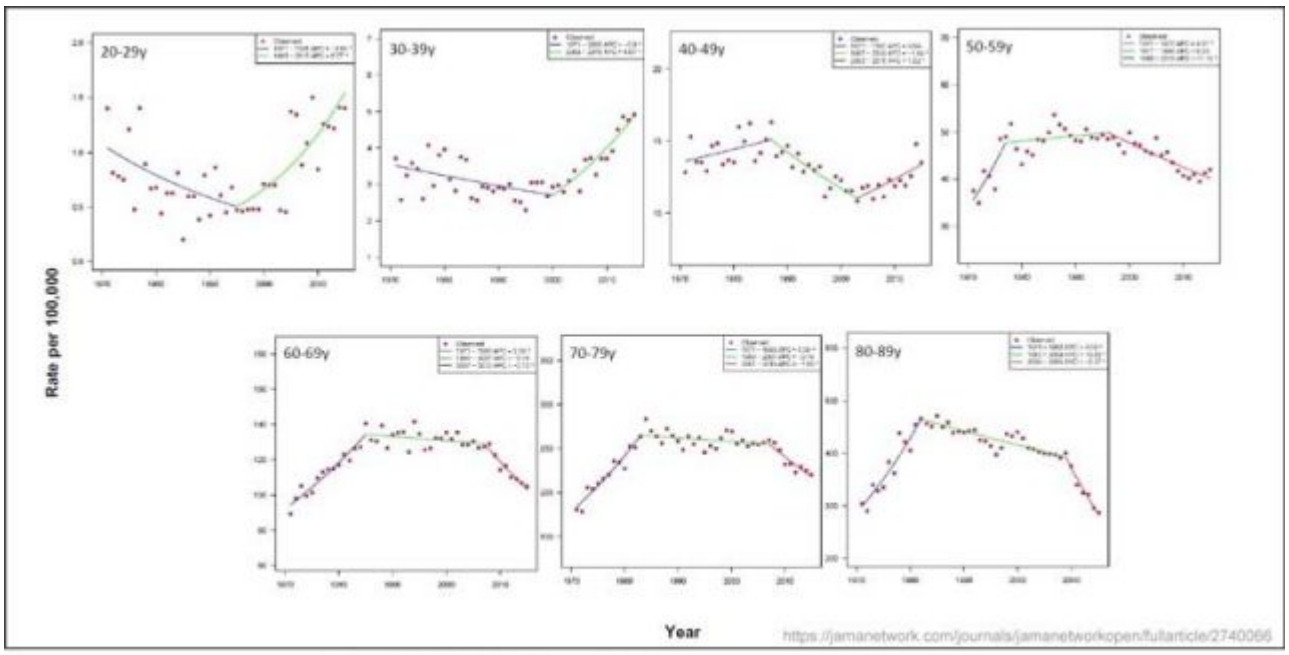
Natürlich erhebt sich da eine Folgefrage: ***Ist das, was sie registriert haben, wirklich eine Messung dessen, über was sie berichten?***

Jüngst kam mir ein Beispiel aus einem anderen Bereich vor Augen, wie intellektuell gefährlich die kognitive Abhängigkeit (fast schon eine Überzeugung) zu Trendlinien für die wissenschaftliche Forschung sein kann. Man erinnere sich, Trendlinien in aktuellen Graphiken sind oftmals berechnet und von Statistik-Softwarepaketen gezeichnet, und die Ergebnisse dieser Software werden viel zu oft als eine Art enthüllter Wahrheit angesehen.

Ich habe nicht den Wunsch, irgendeine Kontroverse loszutreten über das aktuelle Thema einer Studie, welche die folgenden Graphiken produziert hat. Ich habe die genannten Bedingungen in den Graphiken abgekürzt. Man versuche, mir zu folgen, aber nicht um die medizinische Thematik zu verstehen, um die es geht, sondern um die Art und Weise, mit der Trendlinien die Schlussfolgerungen der Forscher beeinflusst haben.

Hier folgt ein großer graphischer Datensatz aus den Begleitinformationen der Studie:

Man beachte, dass dies Darstellungen der *Häufigkeits-Raten* sind, also die Frage „wie viele Fälle dieser Krankheit pro 100.000 Menschen sind berichtet worden?“ – hier gruppiert um 10-Jahre-Altersgruppen. Man hat farbige Trendlinien hinzugefügt, wo sie glauben (Meinung!), dass signifikante Änderungen der Anzahl der Fälle aufgetreten sind.



Age-Specific Incidence Rates [cases per 100,000 population per year] of CRC among Men Canada by 10-year Age Groups

(Einige wichtige Details, über die später noch gesprochen wird, können im Großbild eingesehen werden.)

Wichtige Anmerkung: Die in dieser Studie untersuchte Bedingung ist nichts, was irgendwie von Jahreszeiten oder Jahren abhängig ist wie etwa Grippe-Epidemien. Es ist eine Bedingung, die sich in den meisten Fällen über Jahre hinweg entwickelt, bevor sie entdeckt und besprochen werden kann – etwas, das nur entdeckt werden kann, wenn es beeinträchtigend wird. Es kann auch durch regelmäßige medizinische Untersuchungen entdeckt werden, welche *nur bei älteren Menschen* durchgeführt werden. Also könnte „jährliche Fallzahl“ keine ordentliche Beschreibung dessen sein, was registriert wurde – es ist tatsächlich eine Maßzahl von „jährlichen entdeckten und bekannt gemachten Fällen“ – nicht tatsächlich *Häufigkeit*, was etwas ganz Anderes ist.

In der veröffentlichten Studie erscheint eine kondensierte Version der Graphiken:

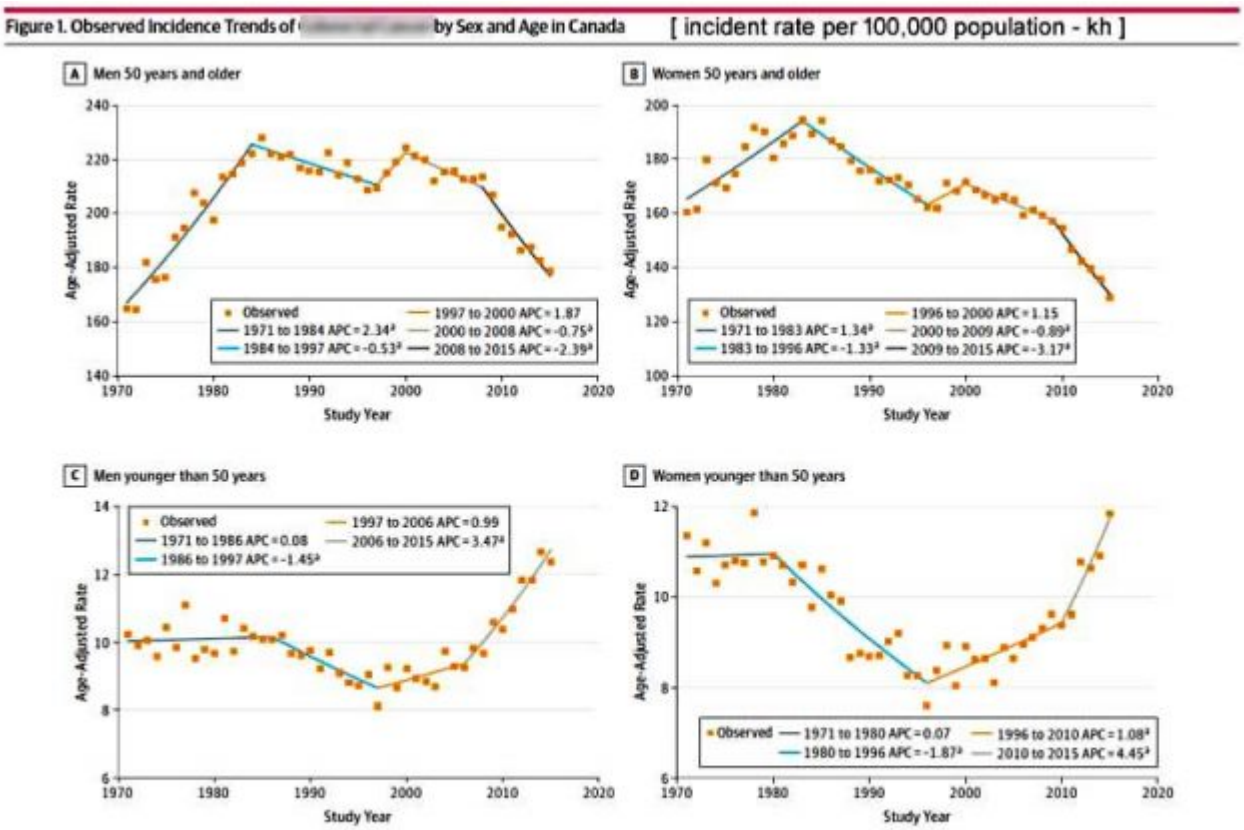


image source: Brenner et al. (2019) --
https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2740066

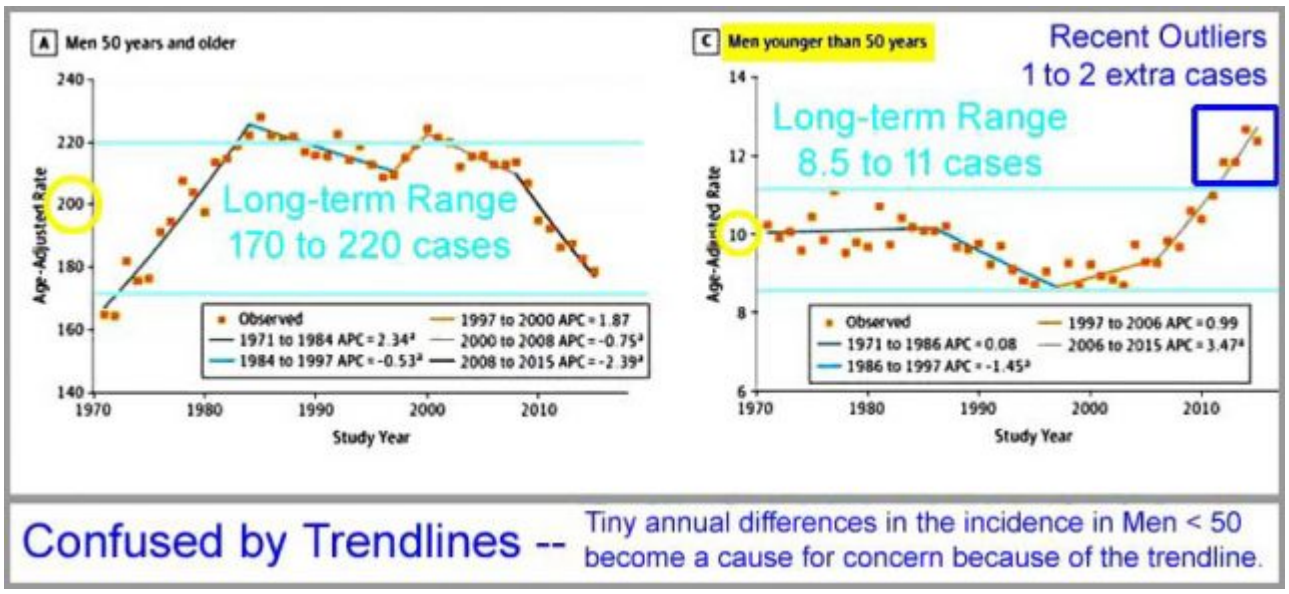
Die älteren Frauen und Männer sind in den Graphiken in der oberen Reihe gezeigt, wobei die Fallzahlen glücklicherweise seit den 1980-er Jahren bis heute rückläufig sind. Durch das mit Bedacht vorgenommene Hinzufügen farbiger Trendlinien steigen die Fallzahlen bei Frauen und Männern jünger als 50 Jahre ziemlich steil. Auf dieser Grundlage (und einer Menge anderer Überlegungen) ziehen die Forscher folgende Schlussfolgerung:

CONCLUSIONS AND RELEVANCE
This study found increasing incidence of CRC diagnoses among Canadian men and women younger than 50 years of age. This increase in incidence among a low-risk population calls for additional research on possible risk factors that may be affecting these younger cohorts. It appears that primary prevention should be the highest priority to reduce the number of younger adults developing CRC in the future.

https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2740066

Schlussfolgerungen und Relevanz: *Diese Studie fand ein zunehmendes Vorkommen von CRC-Diagnosen [= Darmkrebs] unter kanadischen Frauen und Männern unter 50 Jahren. Diese Zunahme der Fallzahlen unter einer* Low Risk-*Bevölkerung verlangt nach zusätzliche Forschungen bzgl. möglicher Risikofaktoren für diese jüngere Bevölkerungsgruppe. Es scheint, dass primäre Prävention höchste Priorität haben sollte, um die Anzahl jüngerer Erwachsener mit Darmkrebs künftig zu reduzieren.*

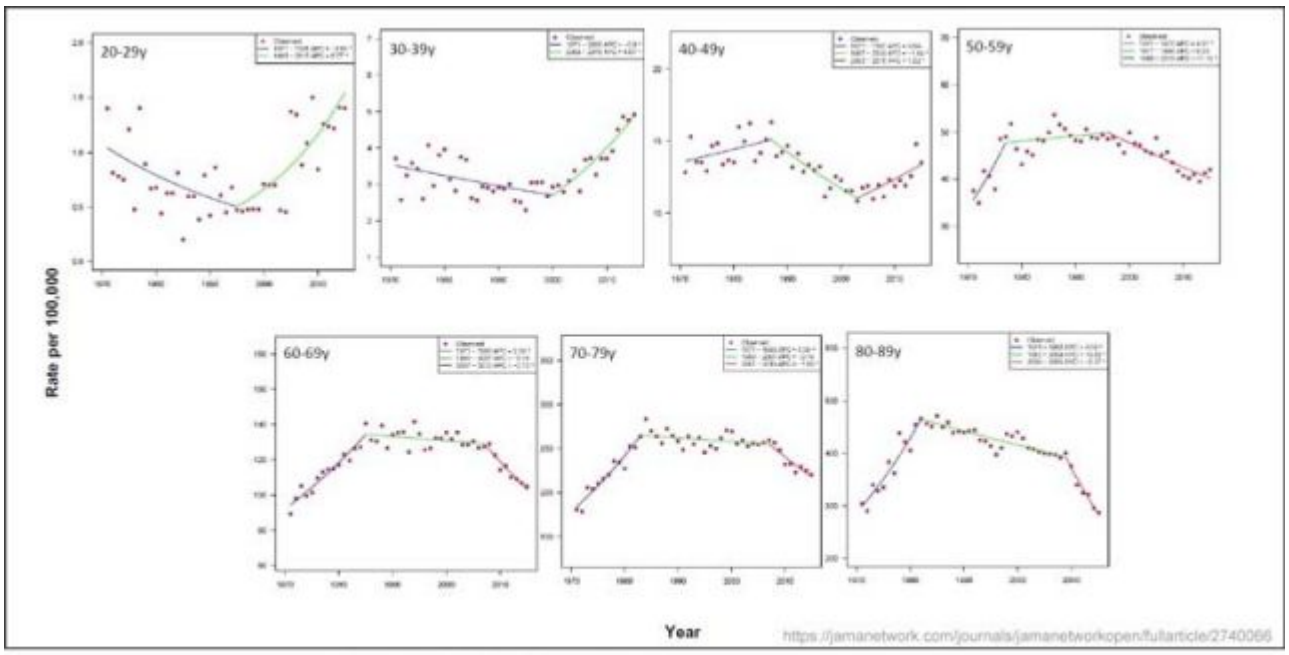
Noch einmal: Es geht mir in keiner Weise um das medizinische Thema hier … sie mögen ja recht haben aus Gründen, die hier keine Rolle spielen. Der Punkt, den ich herüber bringen möchte, ist vielmehr Folgender:



Ich habe zwei der Teilgraphiken mit Kommentaren versehen, in denen es um Fallzahlen bei Männern über 50 Jahre bzw. unter 50 Jahren geht. Über eine Datenlänge von 45 Jahren verläuft die Bandbreite bei Männern älter als 50 Jahre zwischen 170 und 220 Fällen pro Jahr mit einer Varianz von über 50 Fällen pro Jahr. Bei Männern jünger als 50 Jahre lagen die Fallzahlen ziemlich stetig zwischen 8,5 und 11 Fällen pro Jahr pro 100.000 Menschen über einen Zeitraum von 40 Jahren. Erst ganz zuletzt, bei den letzten 4 Datenpunkten, zeigte sich ein Anstieg auf 12 bis 13 Fälle pro 100.000 pro Jahr – eine Zunahme um ein oder zwei Fälle pro Jahr pro 100.000 Menschen. Es kann die Trendlinie für sich sein, die eine Art Signifikanz erzeugt. Für Männer älter als 50 Jahre zeigte sich zwischen 1970 und Anfang der 1980-er Jahre eine Zunahme um 60 Fälle pro 100.000 Menschen. Und doch wird die *entdeckte und berichtete* Zunahme um einen oder zwei Fälle bei Männern unter 50 Jahren als eine Sache „höchster Priorität“ eingestuft – was jedoch in der Realität tatsächlich signifikant sein kann oder auch nicht – und alles könnte sehr gut auch innerhalb der normalen Varianz der Entdeckung und Meldung dieser Art Krankheit liegen.

Die Bandbreite der Fallzahlen bei Männern unter 50 Jahren blieb von Ende der 1970-er Jahre bis Anfang der 2010-er Jahre gleich – das ist ziemlich stabil. Dann gibt es vier etwas höhere Ausreißer hintereinander – mit Zunahmen von 1 oder 2 Fällen pro 100.000 Menschen. Soweit die Daten.

Falls es meine Daten wären – und mein Thema – sagen wir die Anzahl von Monarchfaltern in meinem Garten pro Monat oder so, würde ich aus dem Paneel mit den sieben Graphiken oben entnehmen, dass die Trendlinien alles konfus machen. Hier noch einmal:



Age-Specific Incidence Rates [cases per 100,000 population per year] of CRC among Men Canada by 10-year Age Groups

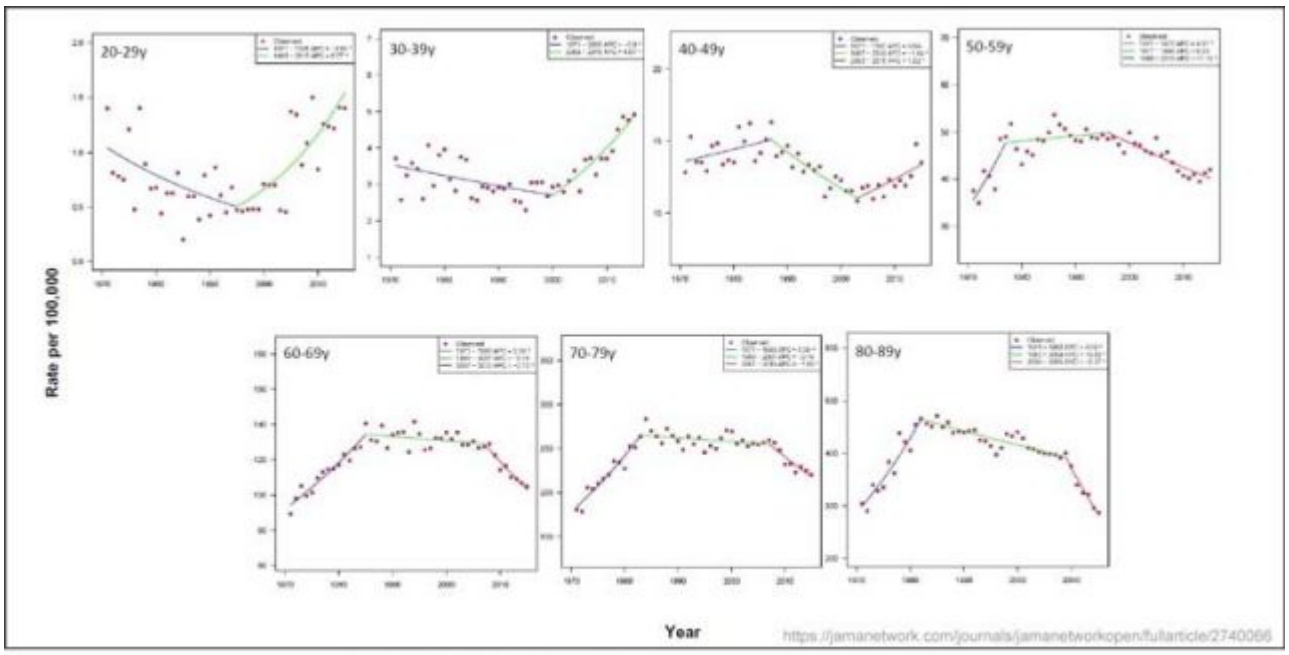
Falls wir mal versuchen, die Trendlinien zu ignorieren, erkennt man in der ersten Teilgraphik, dass die Fallzahlen im Alter von 20 bis 29 Jahren im derzeitigen Jahrzehnt gleich hoch sind wie in den 1970-er Jahren – es gibt keine Änderung. Die Bandbreite liegt hier unter 1,5 Fälle pro Jahr.

Betrachtet man die Untergraphik 40 bis 49 Jahre, erkennt man, dass die Bandbreite etwas gesunken ist, aber die gesamte Größenordnung der Bandbreite beträgt weniger als 5 Fälle pro Jahr pro 100.000 Menschen. In dieser Altersgruppe wurde eine Trendlinie gezogen, welche eine Zunahme während der letzten 12 bis 13 Jahre aufweist, aber die Bandbreite ist gegenwärtig niedriger als während der 1970-er Jahre.

In den übrigen vier Teilgraphiken erkennt man Daten mit der Form eines „Buckels“, welche über 50 Jahre in jeder Altersgruppe die gleiche Bandbreite zeigen.

Es ist wichtig, sich daran zu erinnern, dass es hier nicht um eine Krankheit geht, deren Ursachen bekannt sind oder für die es eine Präventionsmethode gibt, obwohl man behandelt wird, wenn die Krankheit früh genug erkannt wird. Es ist eine Klasse von Krebs, und dessen Vorkommen wird nicht durch Maßnahmen bzgl. der öffentlichen Gesundheit kontrolliert, um diese Krankheit zu verhindern. Derartige Maßnahmen führen nicht zu einer Änderung der Fallzahlen. Man weiß, dass es altersabhängig ist und öfter bei Männern und Frauen auftritt, wenn sie altern.

Es ist diese eine Teilgraphik der Altersgruppe von 30 bis 39 Jahren, welche eine Zunahme der Fallzahlen um 2 Fälle pro Jahr pro 100.000 Menschen zeigt, was den Faktor kontrolliert, durch welchen sich bei Männern jünger als 50 Jahre diese Zunahme zeigt:



Age-Specific Incidence Rates [cases per 100,000 population per year] of CRC among Men Canada by 10-year Age Groups

Man erinnere sich, im Abschnitt *Schlussfolgerung und Relevanz* der Studie wurde dies so beschrieben: „Diese Zunahme der Fallzahlen unter einer Low Risk-Bevölkerung verlangt nach zusätzliche Forschungen bzgl. möglicher Risikofaktoren für diese jüngere Bevölkerungsgruppe. Es scheint, dass primäre Prävention höchste Priorität haben sollte, um die Anzahl jüngerer Erwachsener mit Darmkrebs künftig zu reduzieren“.

In diesem Beitrag geht es nicht um das Vorkommen dieser Krebsart bei verschiedenen Altersgruppen – sondern es geht darum, wie statistische Software Trendlinien über die Daten legt, welche zu Konfusion und möglichen Missverständnissen der Daten selbst führen können. Ich gebe zu, dass es auch möglich ist, Trendlinien aus **rhetorischen** Gründen über die Daten zu legen (mit der Absicht, Eindruck zu schinden) wie im Beispiel von *Climate.gov* (und Millionen anderer Beispiele in allen Bereichen der Wissenschaft).

Unter dem Strich:

1. Trendlinien sind nicht Bestandteil der Daten. Die Daten sind die Daten.

2. Trendlinien sind immer Meinungen und Interpretationen, welche den Daten hinzugefügt werden. Sie sind abhängig von der Definition (Modell, statistische Formel, Software, was auch immer), der man den „Trend“ verpasst. Diese Meinungen und Interpretationen können valide sein oder auch nicht oder auch unsinnig (und alles dazwischen ebenfalls).

3. Trendlinien sind KEINE Beweise – die Daten können Belege sein, sind aber nicht notwendigerweise Belege für das, was behauptet wird (hier).

4. Trends sind keine Ursachen (hier), es sind Auswirkungen. Trends der

Vergangenheit haben nicht die gegenwärtigen Daten erzeugt. Gegenwärtige Daten werden keine zukünftigen Daten erzeugen (hier).

5. Falls die Daten mittels statistischer Software bearbeitet werden müssen, um einen „Trend“ zu bestimmen, dann würde ich vorschlagen, dass man weitere oder andere Forschungen durchführt oder dass die Daten so hohes Rauschen aufweisen, dass ein Trend irrelevant wäre.

6. Berechneten Trends eine „Signifikanz“ zuzuordnen auf der Grundlage eines P-Wertes ist statistisch invalid.

7. Man lege keine Trendlinien in die Graphiken seiner Daten. Falls die Daten nach bestem Wissen valide sind, braucht es keine Trendlinien, um anderen die Daten zu „erklären“.

Link:
https://wattsupwiththat.com/2019/08/06/why-you-shouldnt-draw-trend-lines-on-graphs/

Übersetzt von Chris Frey EIKE