

Mylná představa "mírného" omikronu

Dr. David Glassman, kardiolog, elektrofyziolog, mistr kuchař a obhájce tzv. oxfordské čárky

Rychlost, s jakou lze v pandemii něco poznat, je omezená - omikron nikdy nebyl mírný.

Označení varianty omikron za "mírnou" bylo nezáměrnou chybou, které bylo možné zabránit, kdybychom rozuměli zpožděným časovým řadám.

Varianta omikron se objevila v posledním listopadovém týdnu. V týdnu od 20. prosince byly publikovány dvě studie, které uváděly, že u varianty omikron je menší pravděpodobnost závažných následků. Jedna uváděla, že je to o 2/3 méně pravděpodobné, druhá, že je to o 80 % méně pravděpodobné.

Tím se zrodil narativ o "mírném" omikronu. Tato "dobrá zpráva" se rychle rozšířila a vznikla legenda o tom, že covid-19 se vyvíjí tak, aby byl méně smrtelný - což má být přirozený krok, který nakonec učiní všechny pandemické nemoci. Endemie, jak hlásaly tweety, byla za rohem.

Aby bylo jasno, tato tvrzení byla tehdy nepravdivá a jsou nepravdivá i nyní. Omikron není mírný. Viry se nepotřebují vyvíjet, aby byly méně smrtelné. A endemičnost nemá s mírností nic společného.

Přesto tyto zprávy zaplavily sociální média a konvenční média se snažila nastavit mikrofon každému, kdo byl ochoten nám říct víc.

Pokud je mi známo, výše odkazované studie omikron za mírný výslovně neoznačily. Místo toho suše informovaly o výsledných statistikách. Slovo mírný se objevilo až o mnoho vteřin později a bylo šířeno uživateli sociálních médií.

Ti, kteří na Twitteru psali o tom, že omikron je mírný, pravděpodobně nečekali, jaké to bude mít důsledky. Koneckonců, limit znaků na Twitteru nutí ke stručnosti a "mírný" je znakově úsporný způsob, jak říct strašně moc. Šlo o pohodlnost, nikoli o zlý úmysl. Potíž je v tom, že "mírný" je nespecifické slovo, ke kterému si každý rychle přiřadil vlastní předsudečnou představu. Bylo snadné zahrabat se hlouběji do konfirmačního zkreslení (tendence člověka upřednostňovat ty informace, které podporují jeho vlastní názor).

Ale neopatrné použití slova mírný k popisu omikronu - což byla pravděpodobně jediná věc na Zemi, která se šířila rychleji než samotná nová varianta - nebylo jedinou vadou vyprávění. 22. prosince nebylo možné odhadnout následky viru, který ještě nebyl znám ani čtyři týdny.

Nyní víme, že infekce omikronem má přibližně poloviční pravděpodobnost úmrtí ve srovnání s Deltou (což je méně než 67-80 procentní pokles uváděných v původním "mírném" článku). Velká část zmíněného poklesu je však výsledkem imunity získané očkováním nebo předchozí infekcí. Neočkovaní snášejí nejhorší následky jen o trochu lépe než předchozí vlny variant. Stejně tak děti, které jsou v nepoměrně větší míře neočkované. Opožděné a přetrvávající kognitivní účinky a cévní příhody se vyskytují ve větší míře než kdykoli předtím. Vyprávění o mírném omikronu vyvracejí v lékařských časopisech i v The Atlantic a na sociálních sítích ti samí lidé, kteří spěchali, aby byli "první na Twitteru," kdo napíše o jeho mírnosti. Bude trvat ještě dlouho, než se dozvíme, kolik postižení omikron v naší populaci způsobí. Omikron není mírný podle žádné smysluplné definice - a už teď je brutální ve velkém měřítku - a to ještě nedosáhl toho nejhoršího. Případy možná na mnoha místech dosáhly vrcholu, ale úmrtí a hospitalizace teprve začínají stoupat a BA.2 čeká v záloze.

Tento článek je o čase... a o zpoždění: konkrétně o rychlosti kauzality v pandemii. Jak rychle je možné se o nějaké nové variantě něco dozvědět? Když 22. prosince vyšla první studie označující omikron za mírný, bylo příliš brzy něco o omikronu vědět. Nebylo možné, aby tuto informaci bylo možno z dat získat. S tímto vědomím jsem na Twitteru napsal svou námitku svým 50 sledujícím. To samozřejmě nic nezměnilo.

Škody způsobené prvními zprávami o "mírném" omikronu se projevily jejich následným vlivem na chování jednotlivců i na veřejnou politiku.

Toto nemá být vědecký článek, můj blog k tomu není vhodné fórum. Je to názorový článek. Týká se rychlostní hranice poznání příčinných souvislostí během pandemie. Je určen různorodému publiku - něco si zde najdou lékaři, epidemiologové i všichni ostatní. Omlouvám se, že článek nevyzní víc bombasticky - měl jsem větší plány, ale už takhle je text delší, než jsem zamýšlel.



Tento článek je rozdělen do 8 částí, které jsou zde shrnuty. Kurzívou je uveden hlavní bod každé části. Ačkoli v tomto článku není nikde žádná matematika, některé části jsou docela hutné. Klidně přeskočte některou z částí, ale body, které shrnuji kurzívou, jsou zásadní pro závěr článku.

1. Důležitost nalezení přesného zpoždění: *Na počátku pandemické vlny, kdy počet případů a úmrtí stále roste, vede použití příliš krátkého zpoždění pro stanovení CFR k podhodnocení míry úmrtnosti.*

2. Co znamená pojem "vlna" a co míra smrtnosti v rámci vlny? *Jakýkoli pokus zjistit zpoždění na počátku vlny, kdy jsme ještě na vzestupu, vede k určení zpoždění, které je kratší než skutečné zpoždění.*

3. Doba od stanovení diagnózy do úmrtí: vliv náhodnosti: *Náhoda zkresluje časovou řadu úmrtí. Ale i když se tvar časové řady změní, plocha pod křivkou pro úmrtí musí být se zkreslením stejná jako bez zkreslení. Přítomnost náhodnosti v době od stanovení diagnózy do úmrtí a požadavek kauzality nutně snižuje vrchol, rozšiřuje základnu a přesouvá část plochy křivky na pravou stranu za zenitem křivky úmrtnosti. Jelikož je tato masa na pravé straně křivky, nelze ji včas odhalit.*

4. Zkreslení v druhé jihoafrické vlně: *Demonstrace bodu 2 na případě Jihoafrické republiky (Pokus o zjištění zpoždění na počátku pandemické vlny, když jsme ještě na vzestupu, vede k určení zpoždění, které je kratší než skutečné zpoždění)*

5. Skutečné zpoždění omikronu: *Zpoždění od diagnózy do úmrtí je u omikronu asi o 10 dní delší než u předchozích vln. Ze šumu se zatím vynořily pouze údaje z Jihoafrické republiky. V dalších zemích a městech však začne být brzy jasno.*

Závěrečné myšlenky: *Přestaňte používat slovo „mírný“. Mluvíme o pandemii. Pandemie jsou špatné. Tahle se nijak neliší. Jak moc je omikron zlý? To se dozvíme, až skončí, ale ne dřív.*

Příloha A: *Základní informace o křížové korelaci*

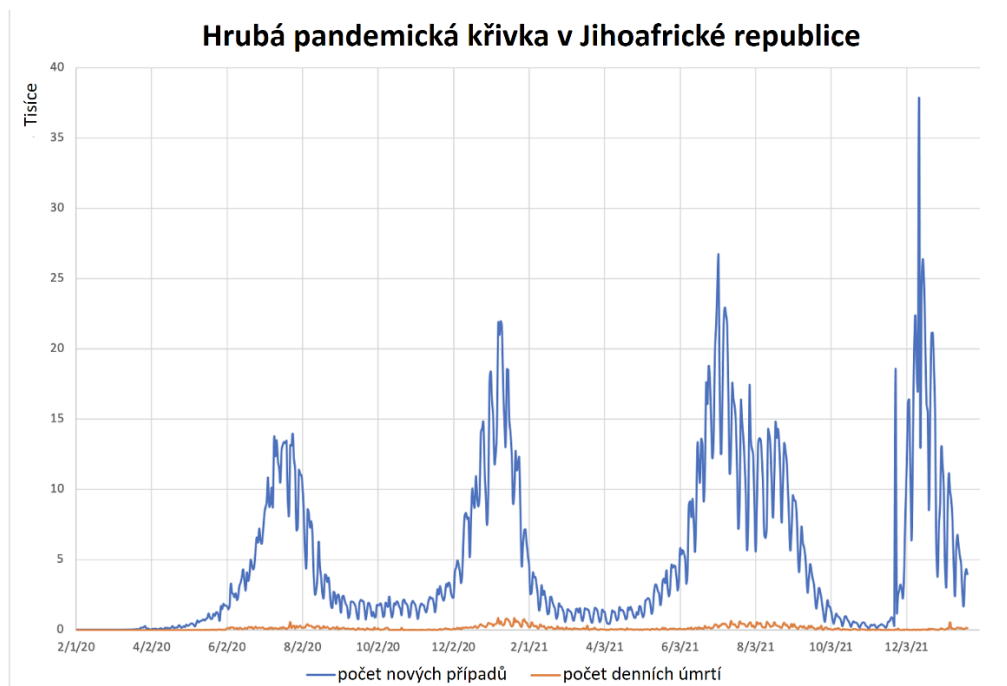
Příloha B: *Křížová korelace na nižších úrovních ve světě.*



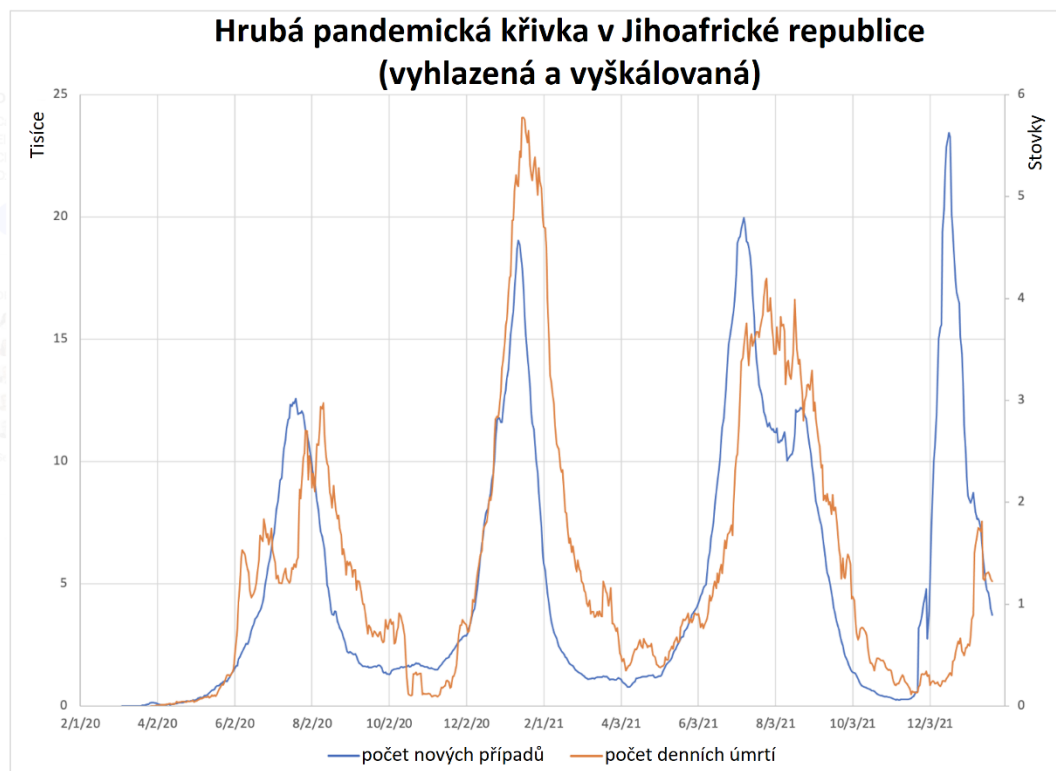
Základní myšlenky a ustálené koncepty použité ve zbytku článku

Časová řada je uspořádaná posloupnost hodnot, které mají nějaký vztah k času: průměrná denní teplota v Třeboni, roční sněhové srážky ve Špindlerově Mlýně, počet lidí procházejících každý týden Hlavním nádražím. V naší současné situaci je relevantní časovou řadou počet denních případů covid-19 a počet úmrtí. Zpožděné časové řady jsou takové, v nichž se jedna časová řada opoždí za druhou, často v důsledku příčinné souvislosti. Měsíční pouštní srážky, růst vegetace a populace hlodavců jsou příklady zpožděných časových řad: nejprve prší, pak rostou rostliny, pak hlodavci přeměňují nové zdroje na potomstvo. Stejně tak úmrtí na covid-19 se opoždí za případy nákazy covid-19.

Následující graf ukazuje téměř dvouletou trajektorii pandemie covid-19 v Jihoafrické republice od 1. února 2020 do 21. ledna 2022.



Tento graf použijí ke zdůraznění některých základních myšlenek a k uvedení několika konceptů, které se objeví ve zbytku článku. Nové diagnózy covid-19 jsou vyznačeny modře, úmrtí oranžově. Jihoafrická republika zažila 4 vlny pandemie. Z těchto nevyhlazených a neškálovaných údajů je těžké vyčíst něco jiného. Artefakty ve výkazech způsobené rozdělením času na 5 pracovních dnů a 2 víkendové dny vytvářejí cyklický artefakt s periodou 7 dnů, který se vyskytuje v časové řadě jak u "případů", tak u "úmrtí". Tato skutečnost v kombinaci se zhruba 50násobným rozdílem v měřítku časových řad zastírá vše ostatní. Pokud však data vyhladíme a změníme jejich měřítko, snadno se projeví zpoždění mezi diagnózou případu a úmrtím.



Zde byly obě časové řady vyhlazeny pomocí klouzavého sedmidenního průměru. Zpoždění mezi případy (modrá) a úmrtími (oranžová) je nyní kvalitativně patrné. Jak bychom však měli toto zpoždění měřit, pokud jej chceme kvantifikovat? Mohlo by být lákavé porovnat vrcholy jednotlivých vln a změřit rozdíl v načasování mezi vrcholem "případů" a vrcholem "úmrtí" pro každou vlnu. Tímto způsobem se však celé vyrovnání vsadí na jediný bod - vrchol. Pokud je poloha tohoto jediného bodu poškozena šumem nebo zkrácením (později uvidíme, že tomu tak nutně je), neexistují žádné další body, které by se používaly ke korekci vyrovnání. Nebylo by lepší, kdybychom mohli nějakým způsobem korelovat více prvků každé časové řady místo jednoho? Přichází křížová korelace!

Křížová korelace je matematická technika z oblasti stochastického (náhodného) zpracování signálů. Používá se ke zjištění zpoždění mezi dvěma zpožděnými časovými řadami. Mezi příklady jejího použití patří audiotechnika, zpracování obrazu, radar a sonar.



Je důležité najít přesné zpoždění

V naší současné situaci chceme vědět, jak často má diagnóza covid-19 za následek smrt. Jedním z klinických způsobů, jak to zjistit pomocí standardní metodologie klinického výzkumu, by bylo sledovat po předem stanovenou dobu podskupinu lidí s diagnózou covid-19 a spočítat, kolik z nich zemřelo. Například pokud bylo 1. prosince 2021 diagnostikováno 10 000 lidí s diagnózou covid-19, mohli byste sledovat náhodný vzorek 1000 z nich po dobu 28 dní. Na konci sledovaného období byste počet zemřelých vydělili 1000 a výsledky byste nahlásili pro míru úmrtí 29. prosince. Použili byste standardní statistické metody k určení intervalu spolehlivosti vašeho měření, který vyplývá ze skutečnosti, že váš vzorek nezahrnoval celou populaci.

Ale co když jste zvolili příliš krátkou dobu sledování? Co kdyby průměrná doba od stanovení diagnózy do úmrtí byla 30 dní? V takovém případě by více než polovina úmrtí nastala až po 28 dnech – tedy až po ukončení studie. Vámi vypočtená míra úmrtnosti by byla příliš nízká.

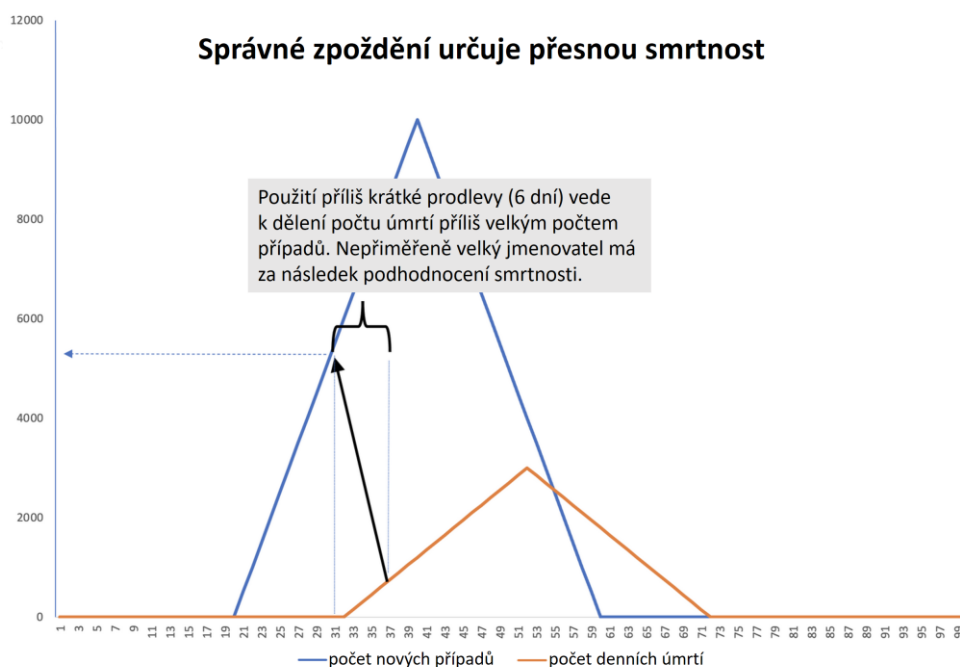
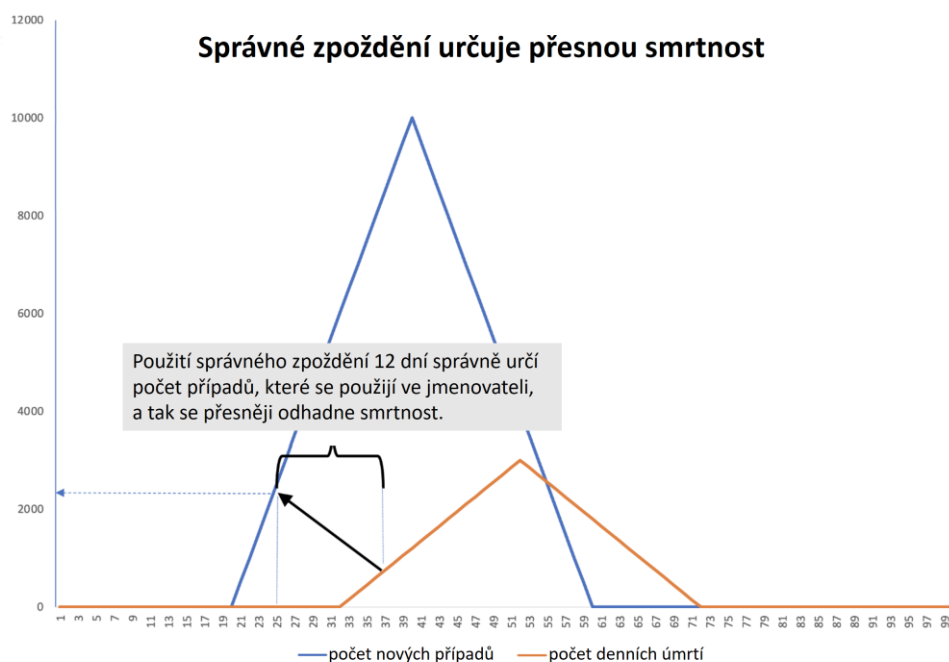
Jinou metodou by mohlo být vydělení všech úmrtí, ke kterým došlo dnes, počtem případů, ke kterým došlo v minulosti. Pokud byste věděli, jaká je průměrná doba od stanovení diagnózy do úmrtí, mohli byste dnešní úmrtí vydělit počtem diagnóz, které se vyskytly v příslušném počtu dní v minulosti.

Představte si, že se pozorováním 3 předchozích vln pandemie určila průměrná doba od stanovení diagnózy do úmrtí na 22 dní. Chcete odhadnout míru úmrtnosti (CFR) současné vlny. Předpokládáte, že dříve stanovená prodleva 22 dní platí i pro tuto vlnu. Dne 23. prosince vydělíte počet úmrtí toho dne počtem případů diagnostikovaných 1. prosince. Ale co když jste zvolili špatné zpoždění? Co když jste měli porovnat dnešní počet úmrtí s počtem případů, které byly diagnostikovány před 30 dny? Nebo před 16?

Nyní prozkoumáme důsledky příliš krátkého odhadu. Pomocí programu Microsoft Excel jsem sestrojil výběrové křivky pro případy a úmrtí. Pro signál "případy" jsem vytvořil trojúhelníkovou vlnu lineárně stoupající od nuly denních případů po 10 000 denních případů v rozpětí 20 dnů. Poté lineárně klesá zpět k nule případů.

Signál "úmrtí" jsem vytvořil na základě předpokladu, že 30 % všech diagnostikovaných zemře přesně 30 dní po stanovení diagnózy. (V pozdější části se budeme zabývat tím, co se stane, když ne všichni zemřou přesně za stejnou dobu po stanovení diagnózy, ale řídí se nějakým rozdělením pravděpodobnosti).

Takto vypadají tyto dva signály.



Začněme prvním grafem. Představte si, že se 37. den rozhodnete vypočítat CFR vydělením počtu úmrtí 37. dne počtem případů, které se vyskytly k nějakému dřívějšímu datu. V den 37 bylo ~700 úmrtí. Předpokládejme, že jste odhadli zpoždění 12 dní. Vrátili byste se k 25. dni a zjistili, kolik diagnóz se ten den vyskytlo. V tomto případě to bylo ~2 100 případů. Dále vydělíte počet úmrtí v den 37 (700) počtem případů v 25. dni (2 100). CFR, které zjistíte, je 33 %, což je velmi blízko hodnotě 30 % naprogramované pro tento příklad.

Nyní si však představte, že jste předpokládali, že zpoždění je pouze 6 dní. Taková situace je znázorněna na druhém grafu. V 37. den bylo stále stejných 700 úmrtí. Když se vrátíte o 6 dní zpět na 31. den, zjistíte, že tehdy to bylo 5 300 případů. Když nyní vydělíte počet úmrtí 37. dne počtem případů v 31. dni, dojdete k závěru, že míra úmrtnosti na případy je 13 %, což je příliš nízká míra. To je obecný výsledek: na počátku

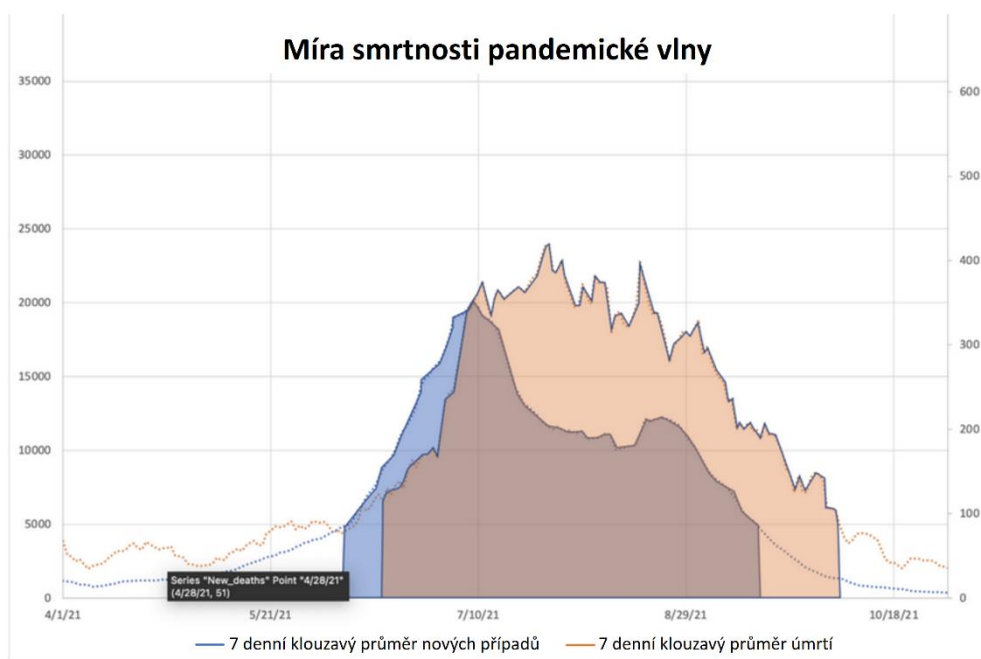
pandemické vlny, kdy případů a úmrtí stále přibývá, vede použití příliš krátkého časového odstavu pro stanovení CFR k podhodnocení míry smrtelnosti.

Nechám na laskavém čtenáři, aby prozkoumal, co se stane, pokud použijete příliš dlouhé zpoždění. Můžete také prozkoumat účinky příliš dlouhého a příliš krátkého zpoždění když vlna pandemie klesá.

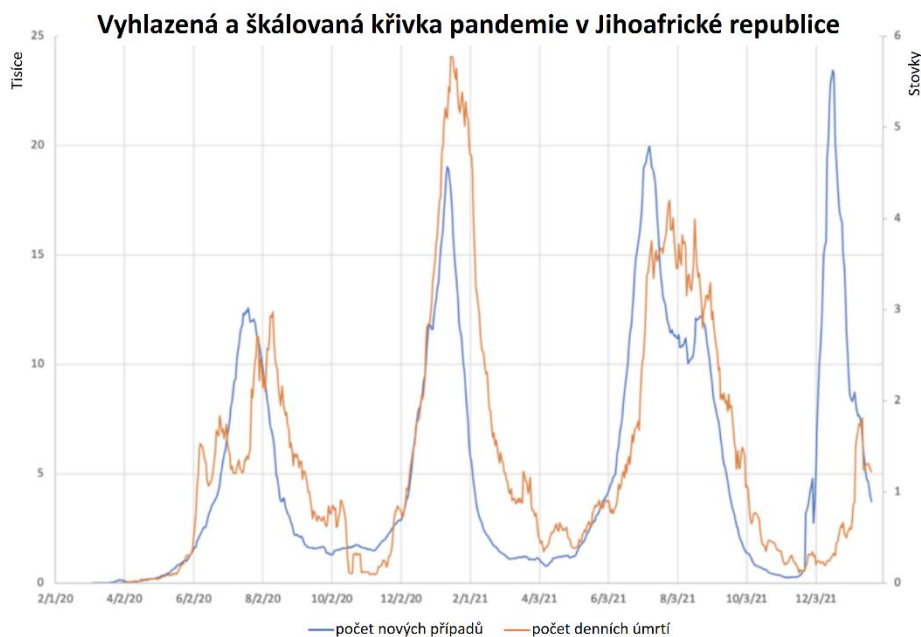


Co je třeba rozumět pod pojmem "vlna" a co mírou smrtelnosti v rámci vlny?

Až tato pandemie dávno skončí, budoucí historici budou konečnou míru smrtelnosti covid-19 uvádět jako podíl celkového počtu úmrtí a celkového počtu případů mezi začátkem a koncem pandemie. Definujme tedy míru smrtelnosti vlny jako plochu pod grafem úmrtí dělenou plochou pod grafem případů - oranžová plocha dělená modrou plochou na grafu níže (velká část těchto dvou ploch se překrývá). Mohli bychom se ptát, zda je tato konečná hodnota zjistitelná dříve, než vlna skončí - například před jejím vrcholem.



Buďme nyní trochu důslednější v tom, co máme na mysli, když mluvíme o "vlně" pandemie. Pravděpodobně tušíte, co to znamená. Kdybych dal následující graf skupině lidí a požádal je, aby zakroužkovali jednotlivé "vlny", pravděpodobně by všichni epidemiologové, senátoři i žáci pátých tříd došli ke stejné odpovědi.



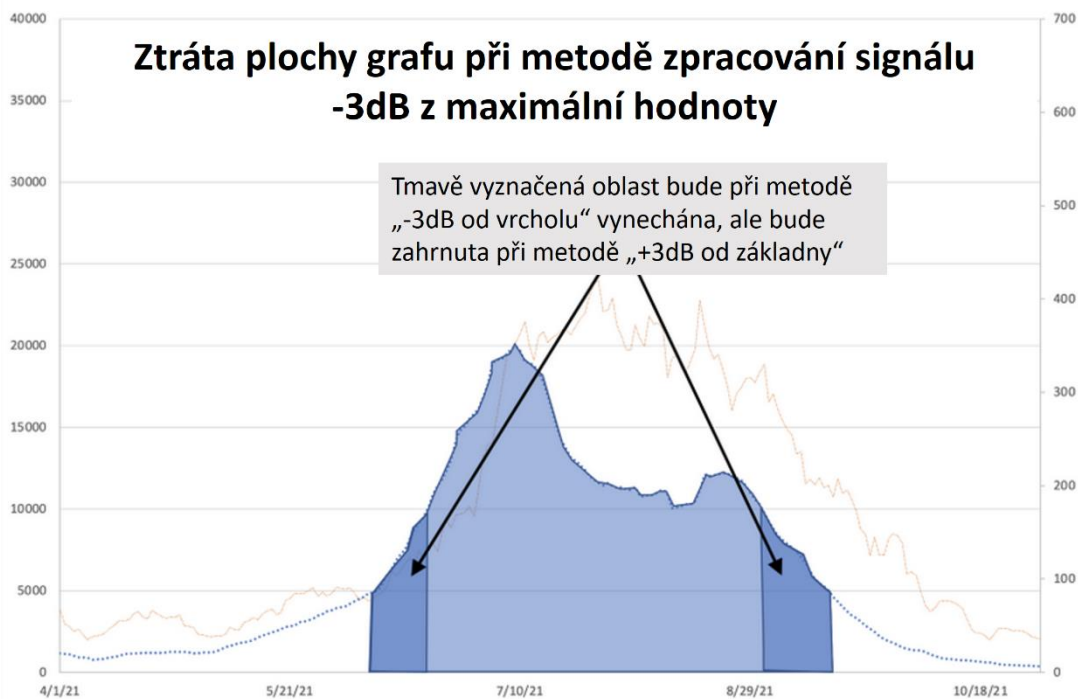
Pokud však chceme začít provádět s těmito daty nějaké výpočty, potřebujeme konkrétnější pojem vlny.

Nedávno jeden tým provedl analýzu popisující tvar průměrné vlny covid-19. Seřadili několik vln pandemie z různých měst a spojili je do prototypové, "průměrné" vlny. Tento obraz průměrné vlny použili k pozorování a vyvození závěrů o stoupající a klesající dynamice vlny. Nyní potřebovali definovat, co tvoří začátek a konec vlny.

K tomu si vypůjčili poznatky ze zpracování signálů. Definovali vlnu jako počátek, když 7 denní klouzavý průměr stoupne nad $1/2$ případné vrcholové hodnoty vlny. Konec vlny definovali jako okamžik, kdy 7 denní klouzavý průměr klesne pod stejnou hodnotu na pravé straně křivky.

Pokud si dovolíme jistou míru libovůle, je tato myšlenka převzata ze standardu používaného při zpracování a filtrování signálů, který je znám jako bod -3 dB (minus tři decibely). Obecný princip této definice se mi líbí, ale pro naše účely má několik omezení. Za prvé, před definováním počátečního bodu musíte vědět, co je vrchol, což znamená, že tuto metodu nelze použít k určení počátku vlny v reálném čase, jak se vyvíjí. Počátek vlny můžeme určit pouze zpětně.

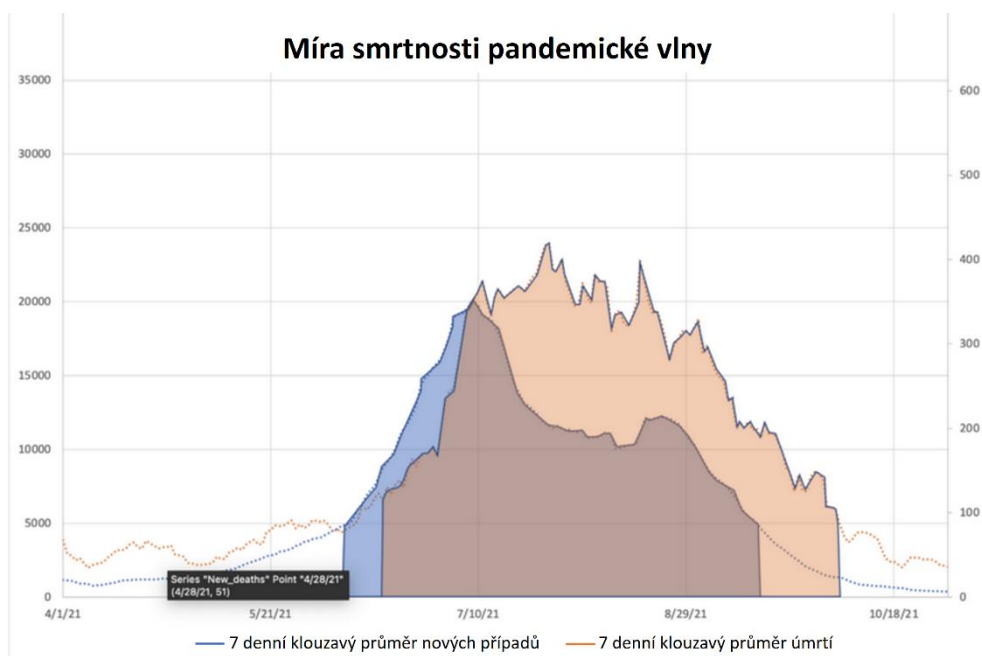
Dalším problémem je, že odříznutí každého konce vlny na úrovni poloviny jejího vrcholu vylučuje velkou část hmoty vlny, kterou pravděpodobně chceme do naší analýzy zahrnout. To je znázorněno na následujícím grafu.



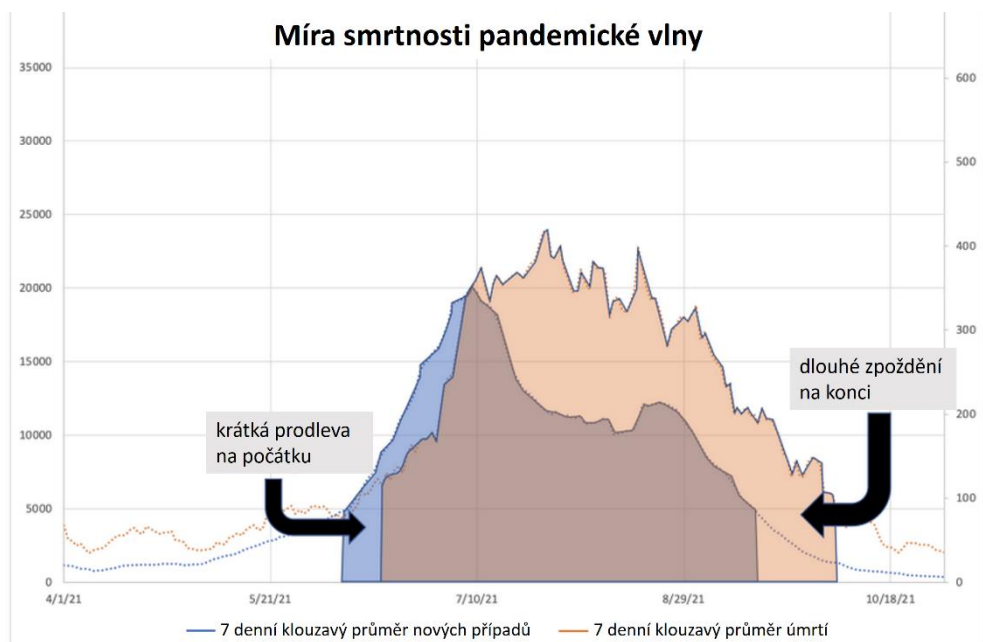
V grafu jsou ukázány jen 7 denní klouzavé průměry nových případů. Úmrtí zde neukazují kvůli přehlednosti, ale koncept platí samozřejmě i pro ně.

Těchto omezení se můžeme zbavit dvěma úpravami. Chceme-li identifikovat novou vlnu v reálném čase při jejím začátku, potřebujeme jinou referenci. Pro účely následující diskuse proto budu definovat začátek vlny jako časový okamžik, kdy se průměrný výkon zdvojnásobí oproti předchozímu minimu (namísto okamžiku, kdy je poloviční oproti nadcházejícímu vrcholu). To nám umožní identifikovat vlnu v okamžiku jejího začátku.

Pro vlnu covid-19 v Jižní Africe, která bezprostředně předcházela vlně omikronu, to vypadá jako na následujícím grafu.



Tady si dobře všimněte prodlevy mezi úmrtími a diagnózou. Skutečnost, že prodleva existuje, je očekávaná, jak jsme již uvedli. Ale trochu neočekávané však je, že zpoždění na konci je delší než zpoždění na začátku. Jinými slovy, zpoždění od konce vlny případů do konce vlny úmrtí je delší než zpoždění od začátku vlny případů do začátku vlny úmrtí. Tahle jednoduchá myšlenka se mi těžko vyjadřuje slovy, raději se podívejte na další graf.



Toto pozorování vyplývá ze skutečnosti, že ne každý, jehož osudem je ve vlně zemřít, zemře po přesně stanovené době od stanovení diagnózy. To znamená, že pokud má 10 % lidí s diagnózou covid-19 zemřít a pokud každý, kdo zemře, zemře přesně 20 dní po stanovení diagnózy, pak bychom očekávali, že časová řada úmrtí bude přesně odpovídat časové řadě případů. Byla by sice v grafu posunuta doprava o 20 dní, ale její obrysy by se shodovaly s grafem nových případů.

Úmrtí však nenásledují diagnózy s přesným zpožděním. Pokud je průměrná doba od diagnózy do úmrtí 20 dní, někteří lidé zemřou po 5 dnech, jiní po 30 dnech atd. - očekáváme, že bude existovat určité statistické rozdělení doby od diagnózy do úmrtí. To způsobuje, že časová řada úmrtí je zkreslená ve srovnání s časovou řadou "případů". To také přesouvá masu časové řady "úmrtí" na zadní stranu řady, za vrchol. Pokud se vám to zdálo těžké pochopit, nezoufejte, budeme se tím zabývat dále v další části.

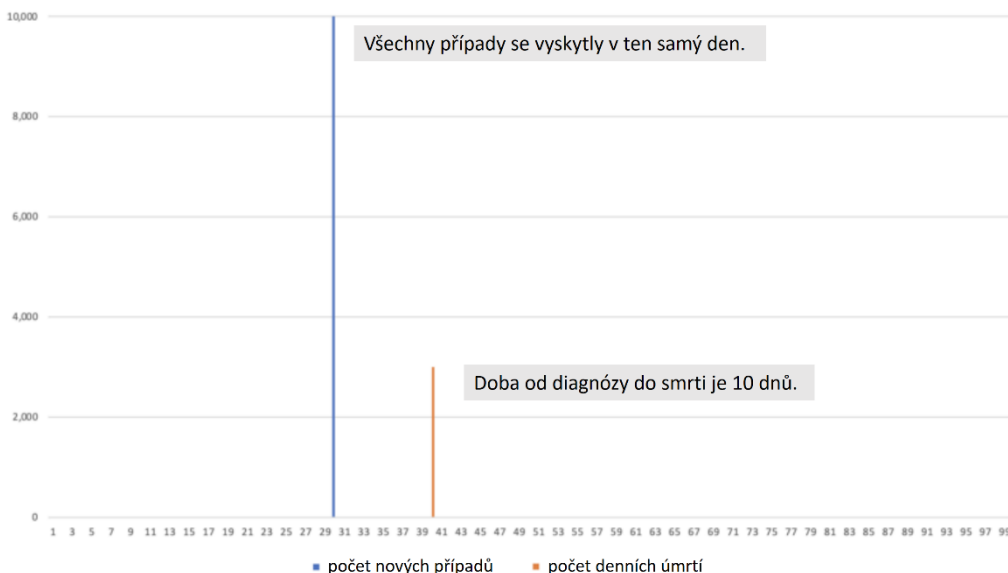


"Doba od diagnózy do smrti": Vliv náhody

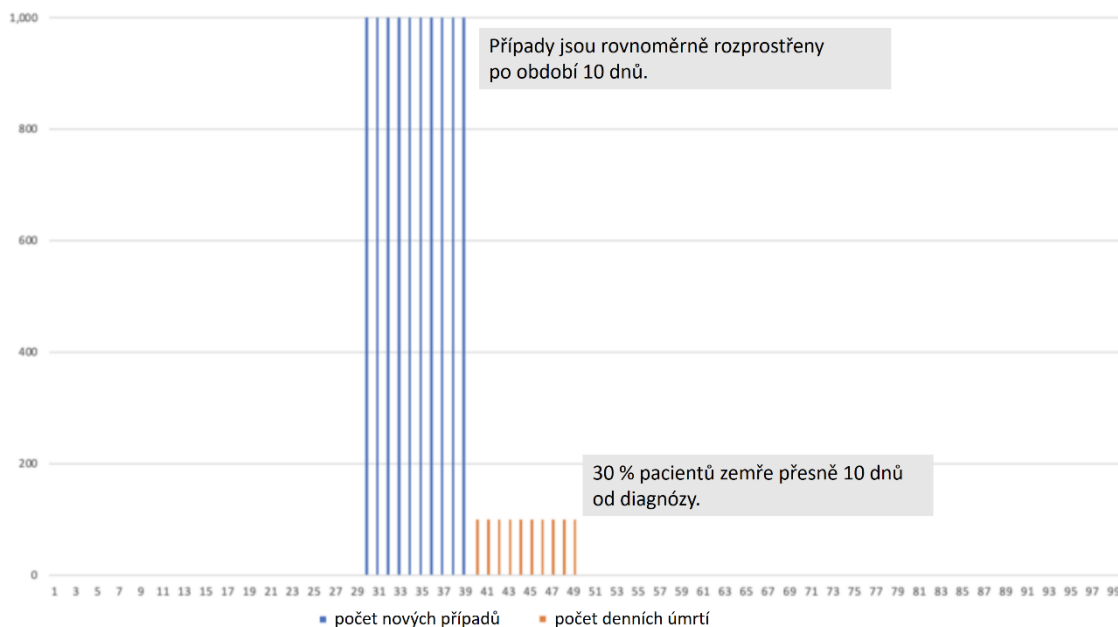
Ne každý, komu je souzeno zemřít na covid-19, zažije přesně stejnou prodlevu od diagnózy do smrti. Někteří lidé mohou zemřít několik dní po stanovení diagnózy, zatímco jiní mohou žít několik týdnů, než své nemoci podlehnou.

Proveďme myšlenkový experiment. Představme si hypotetickou nemoc, kdy každý, kdo onemocní (v tomto případě 10 000 lidí), onemocní přesně ve stejný den. Poté už nikdo další ne onemocní.

Nyní si představme, že úmrtnost je 30 % a prodleva od stanovení diagnózy do úmrtí je přesně 10 dní bez jakýchkoli odchylek. To znamená, že 10 dní poté, co onemocní 10 000 lidí, zemře 3 000 lidí. Tyto grafy by vypadaly následovně:

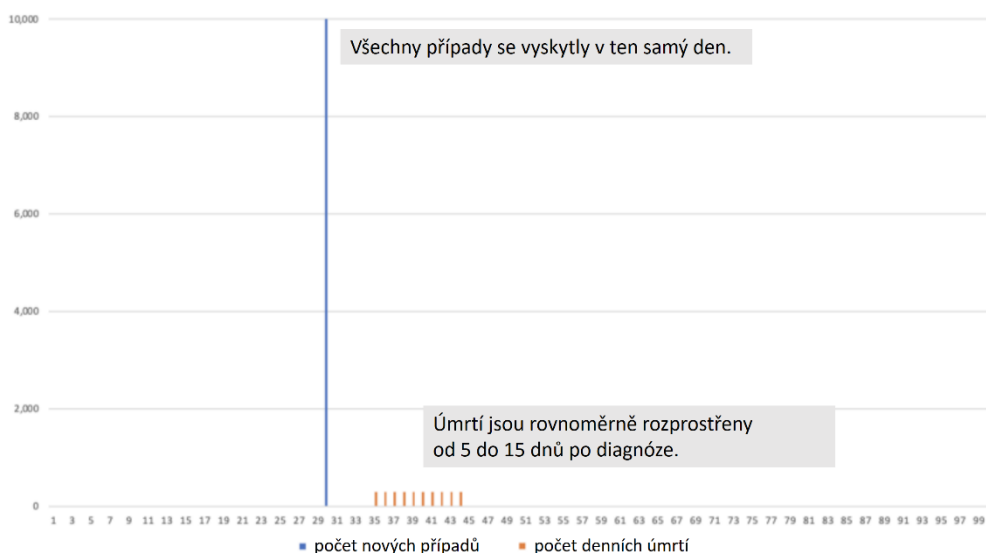


Nyní zmírníme předpoklad, že všechny případy nastaly ve stejný den. Namísto toho předpokládejme, že onemocní stejných 10 000 lidí, ale že onemocní v průběhu 10 dnů, a ne všichni najednou. Smrt ale nastane ve všech případech stále přesně 10 dní po stanovení diagnózy.

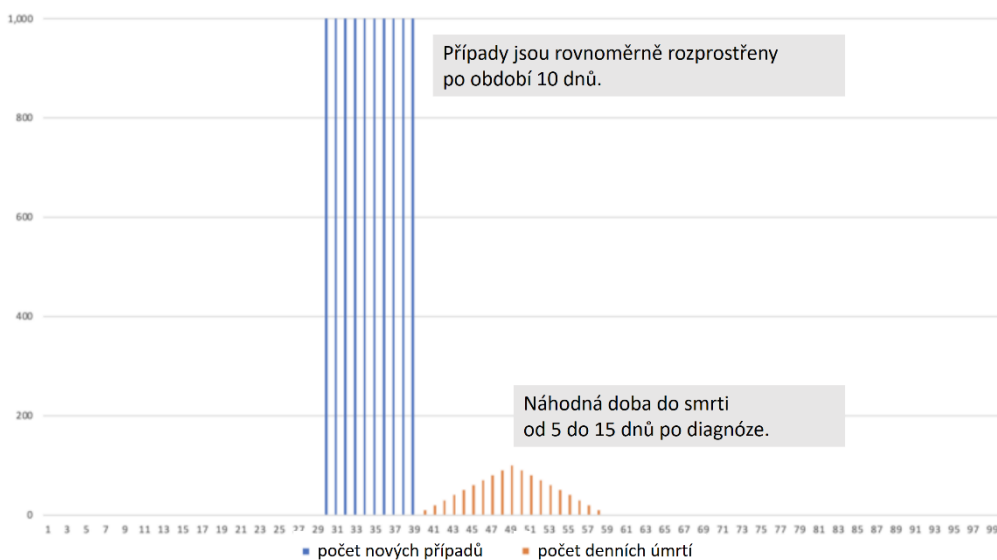


Nyní si představte, že místo pevně stanovené desetidenní doby někteří pacienti zemřou po 5 dnech, jiní po 6 dnech, 7 dnech, 8 dnech atd. Úmrtnost je stále 30 %, takže celkem zemře 3 000 lidí. Tato úmrtí jsou však nyní rozložena. Pokud je doba do smrti rovnoměrně rozložena - řekněme 3% šance na úmrtí každý den od 5. do 14. dne - pak každý den zemře 300 lidí.

Vraťme se ke zjednodušujícímu předpokladu, že všech 10 000 lidí je diagnostikováno ve stejný den. Tento graf vypadá následovně.



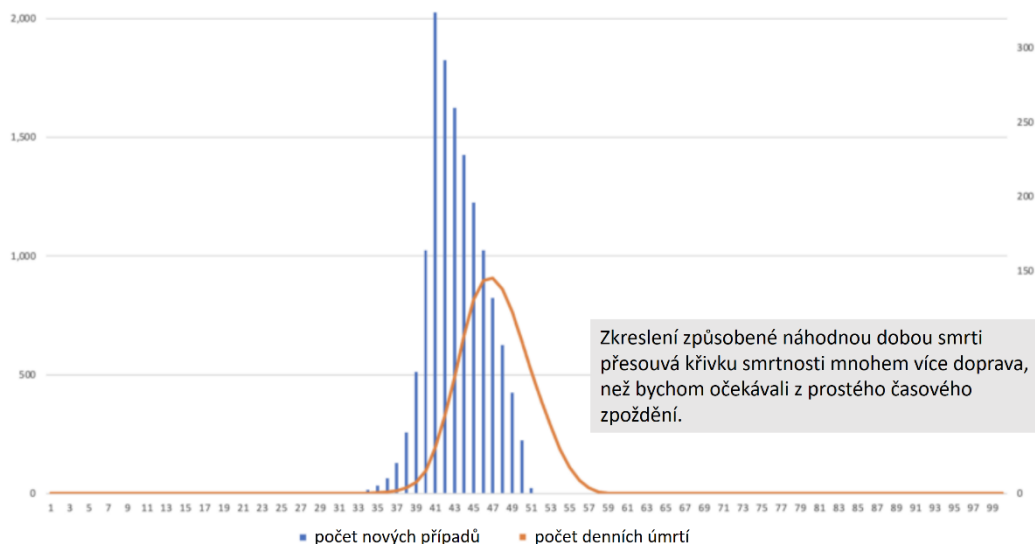
Důsledkem zavedení této nahodilosti je, že úmrtí vzniklá z 10 000 souběžných případů byla v časovém období "rozmazána". Nyní si představme, že diagnózy se opět objevují v průběhu 10 dnů a doba do úmrtí se stále řídí rovnoměrným rozdělením pravděpodobnosti od 5. do 15. dne.



Chtěl bych upozornit na několik věcí týkajících se tohoto grafu. Za prvé, časová řada úmrtí má mnohem širší základnu než časová řada případů. Za druhé, úmrtí nesledují stejný "skok" nahoru k maximální hodnotě a pak "pád" zpět dolů k nule, jako má časová řada případů. Namísto toho postupně (lineárně) stoupá a poté se postupně (rovněž lineárně) vrací na základní hodnotu. Důsledkem této náhodnosti v době do smrti je, že obě řady případů již nemají shodné obrysy. Tento jev se označuje jako "zkreslení".

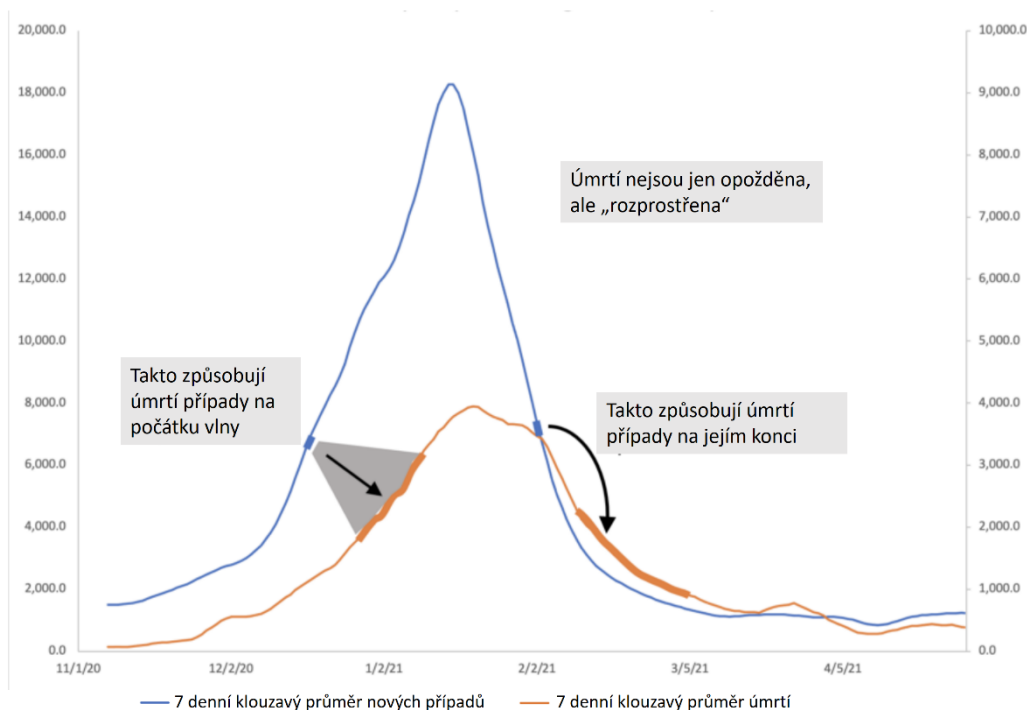
Co je ale důležité - plocha pod křivkou pro úmrtí musí být stejná se zkreslením jako bez zkreslení. Zkreslení nutně snižuje vrchol, rozšiřuje základnu a přesouvá "masu" na pravou stranu křivky smrtelnosti. Zdrojem tohoto zkreslení je nahodilost, která se vyskytuje v různé délce doby od stanovení diagnózy do úmrtí (tj. že ne každý zemře přesně stejný počet dní po stanovení diagnózy).

Skutečné pandemické vlny mají vzestupné svahy, které se řídí exponenciálním růstovým vzorcem, a sestupné svahy, které mají zhruba lineárnější průběh rozpadu. Pravděpodobnostní rozdělení je spíše zhruba Gaussovo. V našem myšlenkovém experimentu by to vedlo k něčemu takovému:

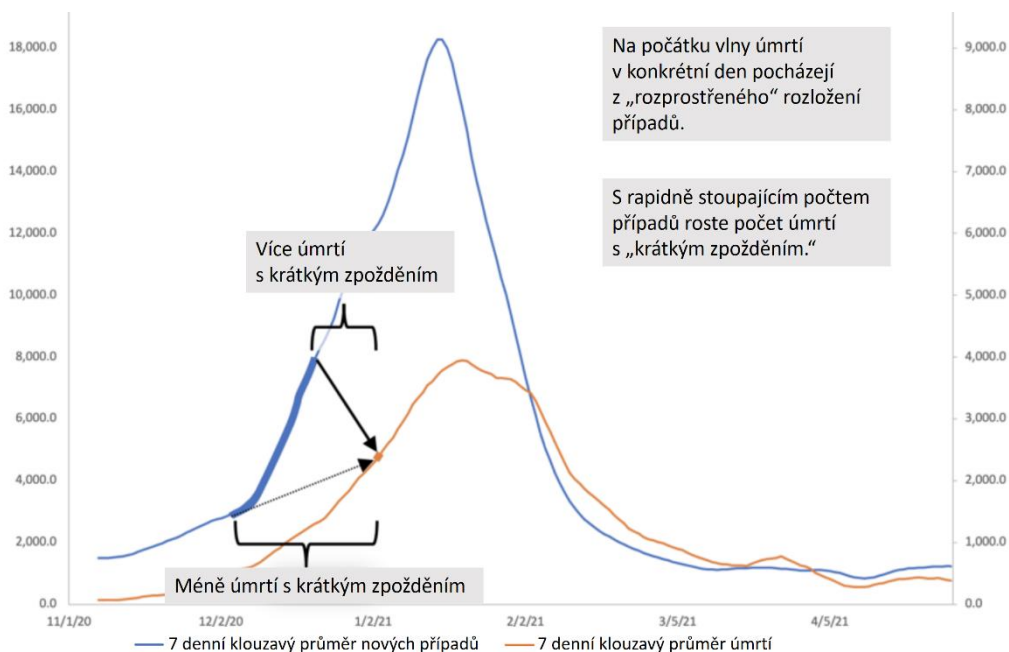


Zkreslení v druhé jihoafrické vlně

Nyní se podívejme na toto zkreslení ve skutečných časových řadách z Jihoafrické republiky. Můžete vidět, jak případy diagnostikované v jeden den vedou k úmrtím podél pásu časové řady pro úmrtí.



A teď provedme stejnou analýzu obráceně. Podívejme se na všechna úmrtí, ke kterým došlo v jeden den, a prozkoumejme, kdy k těmto diagnózám došlo.



Zde vidíte, že úmrtí, ke kterým došlo 5. ledna 2021, byla způsobena diagnózami, které byly stanoveny mezi 8. prosincem 2020 a 29. prosincem 2020. Protože se však nacházíme na vzestupné straně vlny, bylo 29. prosince provedeno mnohem více celkových diagnóz než 8. prosince.

I když je krátké zpoždění 13 dní mnohem méně pravděpodobné než zpoždění 23 dní, díky vysokému počtu diagnóz osoby s 13 denním zpožděním od diagnózy k úmrtí vytvoří klamně velký podíl úmrtí. Kvůli tomu nám pak když zjišťujeme zpoždění na počátku pandemické vlny, nesprávně vychází zpoždění kratší než skutečné zpoždění.



Skutečné zpoždění omikronu

Konečně jsme připraveni podívat se na křížovou korelaci vlny omikron v Jihoafrické republice. Nejprve několik poznámek k terminologii. Nejnovější vlnu v Jižní Africe nazývám vlnou omikron. Uznávám, že (přinejmenším zpočátku) byla tato vlna částečně omikron a částečně delta. Uznávám také, že relativní podíl jedné a druhé varianty se v průběhu této vlny měnil. Uznávám také, že tyto relativní příspěvky se liší nejen v čase, ale také na různých místech světa. Tato analýza se nezabývá vlastní virulencí virové varianty, ale celkovým dopadem všech infekcí během určitého období na populaci - bez ohledu na to, jaké viry mohou cirkulovat. Koneckonců, když se rozhodnete jít ven - s rouškou nebo bez ní, po očkování nebo bez něj - nevíte, s jakou variantou nebo s jakou směsí variant se setkáte.

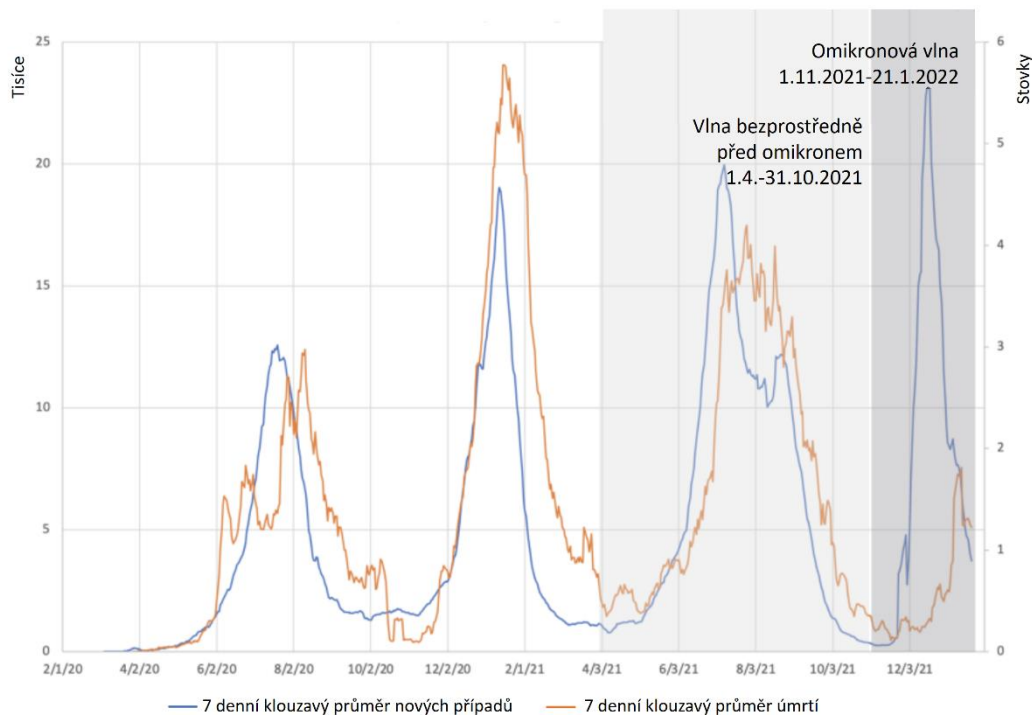
Tato metoda je k této skutečnosti stejně slepá jako my, protože bere v úvahu pouze celý vážený souhrn všech možností. V tomto smyslu tato metoda mnohem přesněji odráží situaci, se kterou se setkáváme při našem rozhodování v reálném světě.

Uvádím pouze jednu zeměpisnou oblast, Jihoafrickou republiku. Je to jediná oblast, jejíž křížový korelační signál k 21. lednu 2022 byl uspokojivý pro vyvození závěrů.

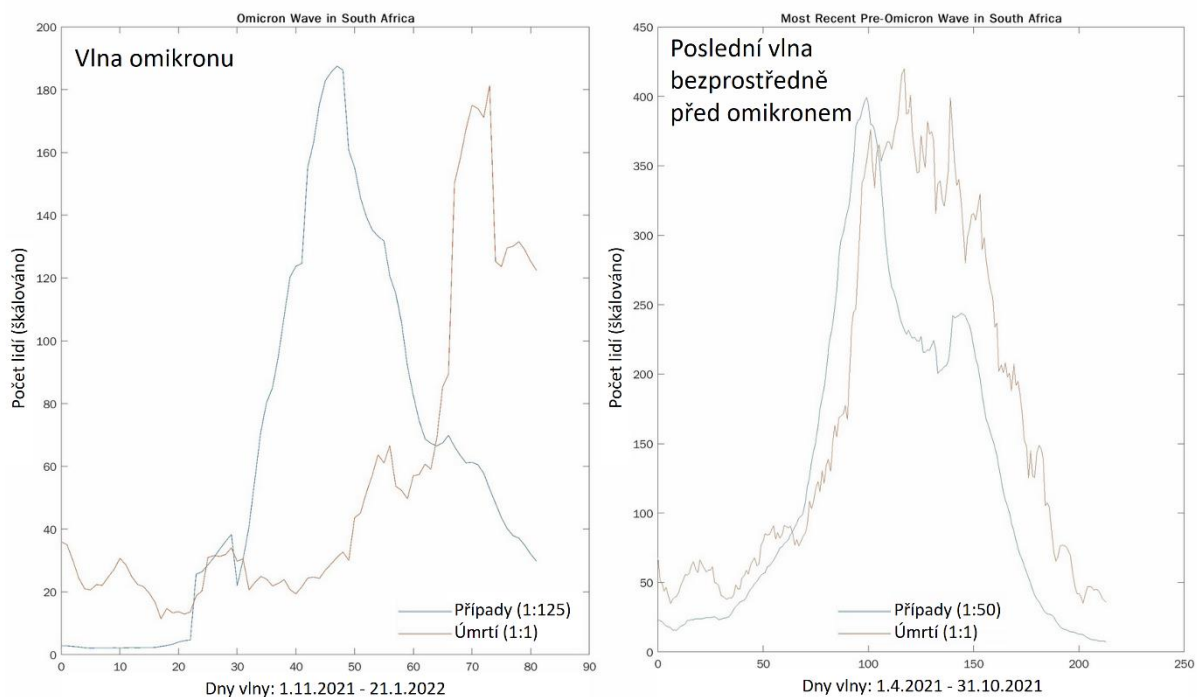
Datové mezníky zvolené pro tyto dvě vlny zahrnují výše popsané počáteční a koncové body "základna +3 dB" plus několik dní datových bodů "navíc", které nazýváme "rukojeti". V případě vlny omikron koncový bod ještě nenastal ani pro jednu z řad.

Soubor dat, který jsem použil, je soubor dat WHO. Časovou řadu jsem vyhledal pomocí sedmidenního průměru, který se běžně používá. Sedmidenní průměr není pro tento účel nevhodnějším filtrem, ale lze jej

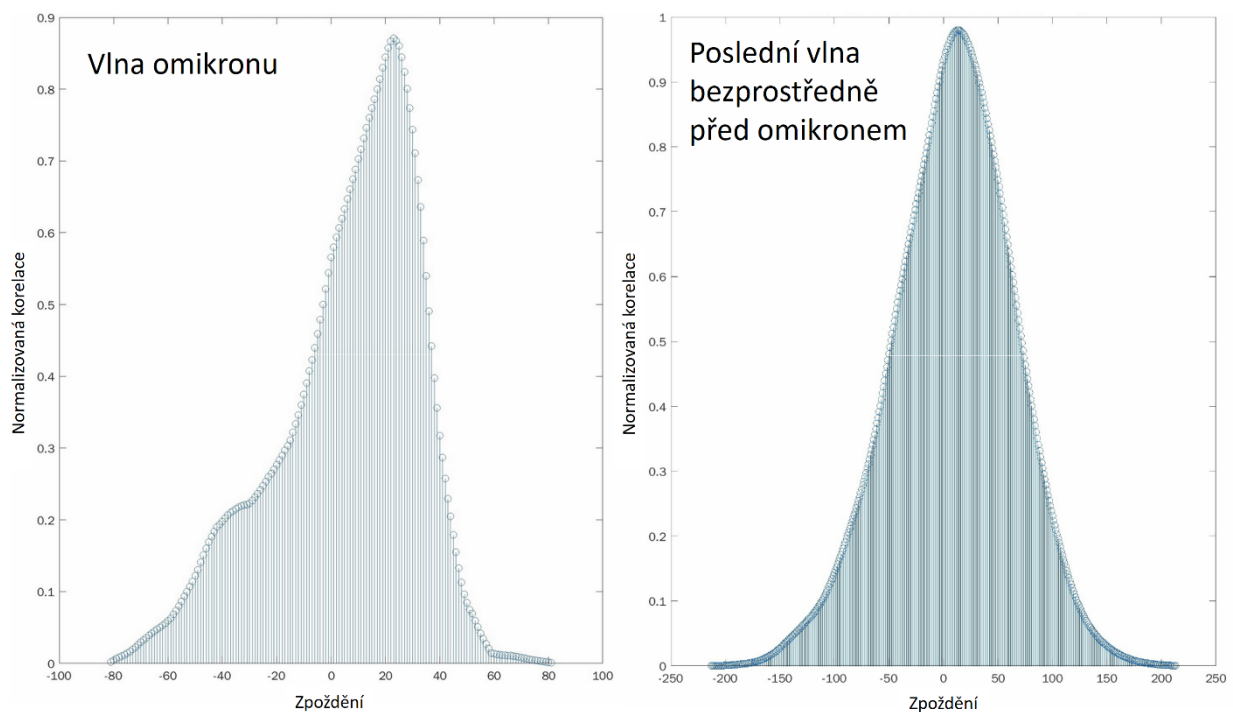
snadno implementovat v aplikaci Excel a umožňuje každému, kdo chce, reprodukovat a ověřit vyhlazenou časovou řadu. Jedná se o 7 denní zpožděné okno (kauzální filtr).



V grafu vidíte čtyři vlny covidu-19 v Jihoafrické republice. Šedě jsou vyznačeny dvě srovnávané vlny, vlna omikron a poslední vlna bezprostředně před omikronem.



Graf uvedený výše ukazuje časové řady pro dvě analyzované vlny. Obě vlny jsem vybral z celkové časové řady.

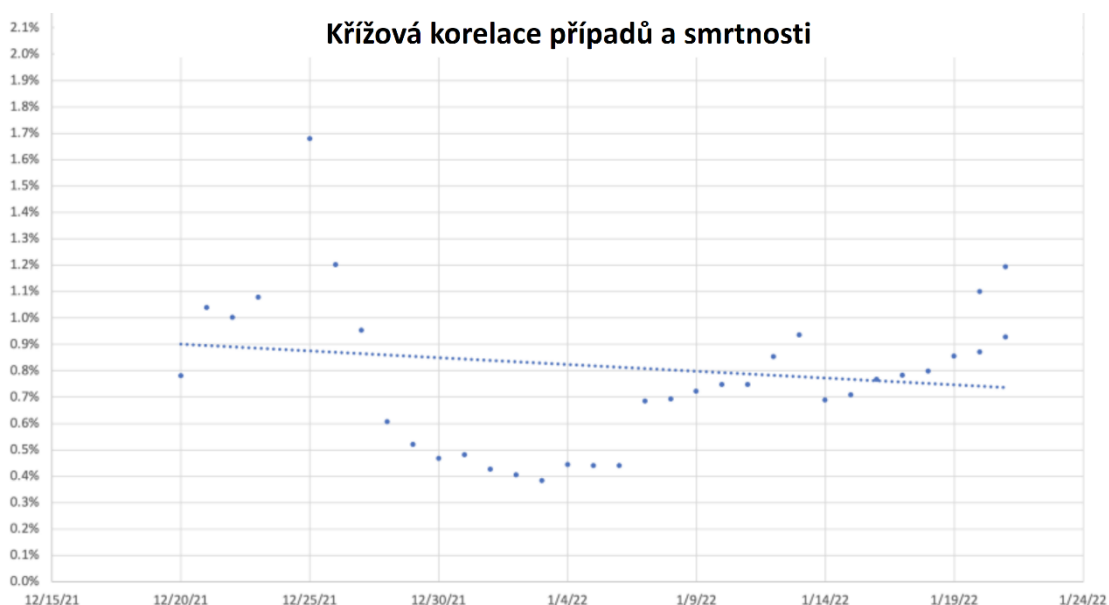


Zde vidíme křížové korelace pro vlnu omikron (vlevo) a poslední vlnu před vlnou omikron (vpravo). Absolutní hodnoty zpoždění při vrcholové korelaci jsou 23 dní a 13 dní. K úmrtí během vlny omikron tedy dochází přibližně o 10 dní později než u předchozí vlny. Zpoždění 23 dní se však blíží zpoždění 22 dní, které Trevor Bedford uvedl u vlny z léta 2020. 13 dní je mnohem méně. Pro toto pozorování nemám vysvětlení. Přesto je zpoždění u současné vlny mnohem delší než u té, která jí bezprostředně předcházela.

Rád bych vás upozornil na další rys křížové korelace. Graf vpravo je hladký a symetrický. Ten vlevo je asymetrický a má podivný "hrb". Ten pochází z toho, že časová řada úmrtí je ve vlně omikron "useknutá", protože ještě nedokončila svůj návrat k základní linii. Již jsem ukázal, že příliš brzké zastavení časové řady zkresluje výsledky a vede k závěru o příliš krátkém zpoždění. Delší čekání na shromáždění více údajů vždy vede k prodloužení zpoždění. Když počkáme, zjistíme, že skutečné zpoždění je o několik dní delší.

Jedním ze způsobů, jak se s tím vypořádat, je odhadnout chybějící "useknuté" hodnoty. Pokud tak učiníte, riskujete, že do svých závěrů vnesete potvrzovací zkreslení. Učinil jsem vědomé rozhodnutí nekategorizovat zkrácené časové řady s "odhadovanými" údaji.

Nakonec použiji zpoždění, které jsme určili, k výpočtu míry úmrtnosti. To je znázorněno na dalším grafu:



Zde jsem vydělil počet úmrtí každý den od 20. prosince 2021 do 21. ledna 2022 počtem případů 23 dní předtím a vypočítal regresi. Očekáváme, že tato přímká bude mít velmi malý sklon, což je tento případ. Upozorňuji na několik odlehklých datových bodů, které se vyskytly v blízkosti Vánoc a mohou být způsobeny artefakty plynoucími ze svátečního volna. Dostáváme výsledek, že CFR je přibližně 0,75-0,9 %, což je o 50 % méně než v předchozí vlně (ale ne 67-80% snížení z prvních zpráv) a blíží se tomu, co jsem odhadoval ve svém článku z 31. prosince 2021 (tedy 45 % snížení).

Závěrečné myšlenky

Je přirozené, že chceme vědět, jak špatná bude nová vlna, už tehdy, když začíná. To však není možné zjistit dříve, než projde hlavní masa vlny. Abychom mohli s jistotou posoudit míru výskytu závažných následků, jako je smrtnost, musí vrchol případů již dávno pominout. Tehdy většina lidí, kteří mají zemřít, již zemřela nebo už alespoň byla vystavena nákaze, která má za následek smrt. Z tohoto důvodu bychom nikdy neměli činit osobní nebo politická rozhodnutí na základě něčího hodnocení závažnosti vlny předtím, než tato vlna téměř skončí. Pokud se mě někdo na začátku příští vlny zeptá: "Jak zlá bude tahle?", odpovím mu: "Řeknu vám to, až skončí."



V překladu jsou vynechány přílohy a technické poznámky. Celý článek v originálu si najdete [zde](#).