

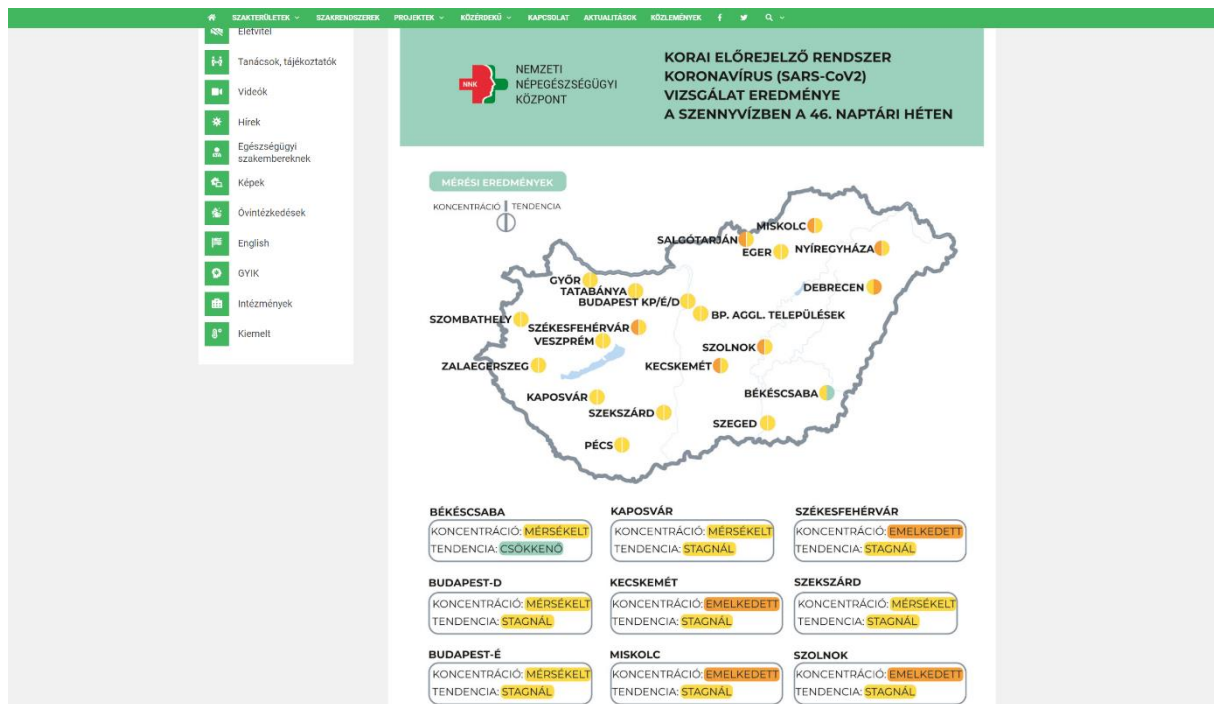
2020. november 16.

A koronavírus járvány lefutásának előrejelzése szennyvízadatok elemzése alapján

A Nemzeti Népegészségügyi Központ (NNK, <https://www.nnk.gov.hu/>) 2020. július 29-e (30. hét) óta 18 város és Budapest három területe (Észak, Dél, Központ) esetében közöl információkat a COVID-19 vírus örökítő anyagának a szennyvízben mért koncentrációjáról. (Az ezt megelőző hetekről is közöl adatokat, de csak kevesebb városról.) A szennyvíz megfigyelése a nemzetközi kutatások szerint jó előrejelzője annak, hogy a fertőzöttség mennyire terjedt el a lakosság körében (lásd például [itt](#), [itt](#)).

Az NNK sajnos *nem közli az eredeti mérési adatokat*, helyette két szempont szerint (koncentráció és tendencia) négy-négy kategóriába sorolja az eredeti információkat és *csak ezeket a szöveges értékeléseket teszi nyilvánossá honlapján* (lásd 1. ábra). A koncentráció esetében: „alacsony”, „mérsékelt”, „emelkedett”, „magas”; a tendencia esetében: „csökkenő”, „stagnál”, „emelkedik”, „erősen emelkedik” szöveges értékeléseket adja meg az általa elvégzett vizsgálatok eredményei alapján ([itt](#)). Ezek a kvalitatív értékelések közvetlenül nem alkalmasak statisztikai elemzésre.

1. ábra: A koronavírus szennyvízben mért koncentrációjának adatai az NNK honlapján, 46. hét (részlet)



Forrás: <https://bit.ly/3nmW7Eu>

Nem tudjuk, mi okból teszi ezt az NNK. Talán a nagy mintavételi hiba miatt tartózkodik az eredeti adatok nyilvánosságra hozásától? Vagy saját adataiban sem biztos? Olyan számítási, egyszerű összeadási hibáktól tart, amelyet például ma követett el a járványüggyel foglalkozó magyar kormányzati apparátus?¹ Vagy azt kívánja elérni a pontos adatok eltitkolásával, hogy minél kevesebb elemzés szülessen ebben a témában? Erre nem tudjuk a feleletet, csak reménykedni lehet, hogy egyik előző feltevés sem igaz.

Bár szöveges (kvalitatív) információkat közöl csak az NNK a magyar állampolgárokkal, de nem lehetetlen e kvalitatív információkat kvantifikálni, feltéve, ha az NNK döntése, hogy az eredeti adatok alapján milyen minősítést („alacsony”, „mérsékelt”, stb.) ír, mindvégig konzisztens. Azaz sem hétről hétre, sem pedig városról városra nem változik az, hogy mit tekint például „alacsony”-nak, sem az, hogy mire mondja azt, például, hogy „emelkedik”.

A szöveges válaszok numerikussá és ezzel elemezhetővé alakítása nem lehetetlen. Carlson és Parkin például a lakosság inflációs várakozásait felmérő közvélemény kutatások kvalitatív válaszainak kvantifikálására dolgoztak egy módszert (lásd [itt](#), és [itt](#), valamint [itt](#)). Az NNK adatainál ez nem alkalmazható, de egy egyszerűbb megoldás már igen. Az empirikus konjunktúra kutatásokban jól ismert „*balance statistics*” (lásd például [itt](#): 89-90. oldal) mintájára át lehet alakítani az NNK által megadott kvalitatív (szöveges) válaszokat úgy, hogy az átalakítás után kapott mutató konzisztens és statisztikai elemzésre is alkalmas legyen.

Ennek érdekében képeztünk egy kompozit indexet (J), amely a koncentrációra (K) és a tendenciára (T) vonatkozóan az NNK által megadott információkat tartalmazza az alábbiak szerint:

$$K_{i,t} = \begin{array}{l} 1 - \text{ha „alacsony”} \\ 2 - \text{ha „mérsékelt”} \\ 3 - \text{ha „emelkedett”} \\ 4 - \text{ha „magas”} \end{array}$$

minősítést adott az NNK az általa mért, de nyilvánosságra nem hozott eredeti *koncentrációs mérési adatok*nak az i városnál a t héten, és

$$T_{i,t} = \begin{array}{l} 1 - \text{ha „csökkenő”} \\ 2 - \text{ha „stagnál”} \\ 3 - \text{ha „emelkedik”} \\ 4 - \text{ha „erősen emelkedik”} \end{array}$$

minősítést adott az NNK az általa mért, de nyilvánosságra nem hozott eredeti *tendencia mérési adatok*nak az i városnál a t héten;

¹ Az NNK adatai alapján 2020. november 16-án 8:50 perckor az jelent meg a <https://koronavirus.gov.hu/> oldalon, hogy „6495 újabb magyar állampolgárnál mutatták ki az új koronavírus-fertőzést (COVID-19), ezzel 147961 főre nőtt a hazánkban beazonosított fertőzöttek száma.” (<https://web.archive.org/web/20201116080528/https://koronavirus.gov.hu/cikkek/6495-fovel-emelkedett-beazonosított-fertozottek-szama-es-elhunyt-93-beteg>) Ezzel szemben egy nappal korábban 9:18 perckor ezt írták „4238 újabb magyar állampolgárnál mutatták ki az új koronavírus-fertőzést (COVID-19), ezzel 140961 főre nőtt a hazánkban beazonosított fertőzöttek száma. Elhunyt 107 többségében idős, krónikus beteg, így az elhunytak száma 3097 főre emelkedett, 31 599-en pedig már meggyógyultak.” (<https://koronavirus.gov.hu/cikkek/4238-fovel-emelkedett-beazonosított-fertozottek-szama-es-elhunyt-107-beteg>). A beazonosított fertőzöttek számának egy napos növekedése megkapható, ha a november 16-án közölt adatból kivonjuk az előző napit, azaz $147.961 - 140.916 = 7000$ fő, ami pedig nem egyezik a 6495 fővel. Erre a hibára több független online portál is felhívta a magyar kormány figyelmét ([itt](#) és [itt](#)). Ezt a hibát később a magyar kormány tájékoztató oldala javította anélkül, hogy a hibára, és ennek javítására külön felhívta volna a figyelmet ([itt](#)).

valamint a kompozit mutató városi és átlagolt értéke

$$J_{i,t} = K_{i,t} + T_{i,t}$$

$$J_t = \frac{\sum_{i=1}^{21} J_{i,t}}{42}$$

minden i -re és t -re.

A fentieknek megfelelően $1 \leq J \leq 4$.

A J minimális értékét ($J=1$) akkor érné el, ha az NNK minden megfigyelési helyen a koncentrációra az „alacsony” és a tendenciára a „csökkenő” minősítést adná, és maximális értékét ($J=4$) akkor kapná, ha az NNK minden megfigyelési helyen a koncentrációra a „magas” és a tendenciára az „erősen emelkedik” minősítést adná. Természetesen helyesebb lenne a $J_{i,t}$ számításánál súlyozott adatokkal (például városonként a lakosság szám szerinti súlyokkal) számolni. Ezt most nem tesszük meg, mivel a célunk csupán annak bemutatása, hogy még az olyan korlátozott információk is, amelyeket az NNK nyilvánosságra hoz, módot adhatnak statisztikai szempontból vizsgálható következtetések levonására.

Ezen túl kiszámítható a J_t szórása is minden t hétre, amely arról ad információt, hogy mennyire terjedt el országosan a járvány. Elképzelhető ugyanis, hogy egyes területeken, városokban alig van jelen a járvány, míg máshol már a járvány nagymértékű elterjedéséről beszélhetünk, de az is, hogy mindenhol egyaránt jelen lévő már a járvány. Egy ilyen számítás is fontos lehet, ha egyébként mód van lokális ellenlépésekre: területenként, városonként eltérő járványügyi intézkedések alapozhatók meg a területi adatok alapján.

Az általunk megadott módszert alapul véve kiszámítható továbbá minden megfigyelt városra vonatkozóan a J mutató értéke, amely jól jelzi, hogy az adott városban milyen lesz a járvány lefutása a következő hetekben.

A 2. ábrán a J mutató értékét úgy ábrázoltuk, hogy a 30. heti (július 20-26.) közötti mérési adatokat vettük 100-nak és ezen érték százalékában ábrázoltuk az ezt követő hetek értékeit. Az ábrán feltehetően a kormányzat november 10-i intézkedéseinek hatása is látszik: a 2020 november 9-15-i héten feltehetően csökkent a szennyvízben a vírus örökítő anyagának koncentrációja Magyarországon.

A 3. ábrán a járvány halálos áldozatainak hetenkénti számát is feltüntettük, mivel ma Magyarországon ezt tekinthetjük a másik érvényes – a járvány lefutását valamennyire jelző – mutatónak. A „beazonosított fertőzöttek” napi száma nem megfelelő erre, mert nagymértékben függ egy külső tényezőtől: hány tesztre került sor az adott napon.

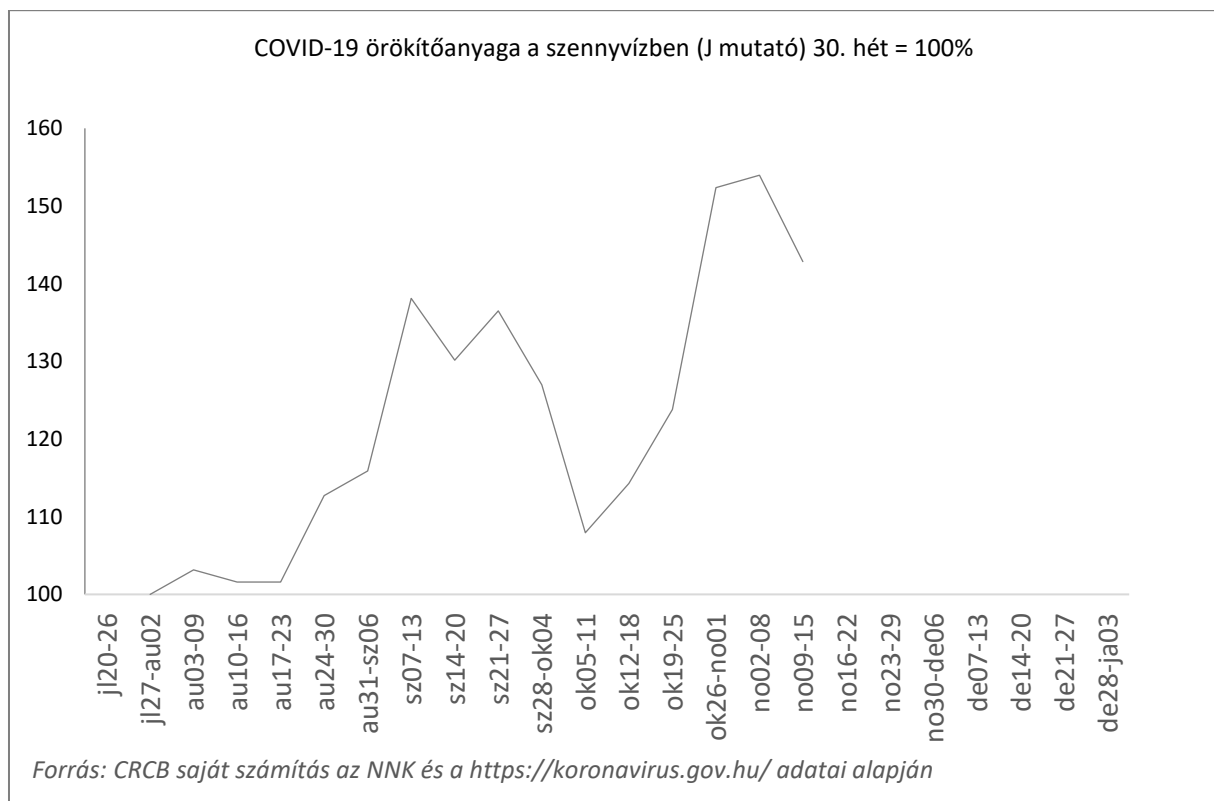
A 3. ábrán a J mutató értékét hat héttel előre helyeztük ($t+6$ hét): vagyis a vírus örökítő anyagának szennyvízben való koncentrációját hat héttel későbbi halálozási adatokkal vetettük össze (így a legerősebb a korreláció a két idősor adatai között). A halálos áldozatok november 9-15-i héten megfigyelhető csökkenő növekedési üteme ezek szerint még nem a kormányzat november 10-i intézkedéseinek, hanem a járvány jóval ezelőtti - szeptember 21-e és október 11-e között megfigyelhető – relatív lanyhulásának köszönhető. Minden egyéb feltételt változatlanul véve, ezek szerint a halálos áldozatok száma november 22-e után újra magasabb ütemben fog nőni. A halálozási adatokban a kormányzati intézkedések hatása csak december második felében lesz érzékelhető.

A 3. ábra csupán illusztráció, mivel nagyon kevés megfigyelés áll rendelkezésre, és az idősor rövidege nem tesz lehetővé semmilyen idősoros elemzést, de arra felhívja a figyelmet, hogy érdemes a kormányzati intézkedések időzítésénél és az intézkedések megtervezésénél tekintetbe venni a szennyvíz elemzéséből kinyerhető információkat. Érdemes lenne nyilvánosságra hozni az eredeti

adatokat is, és érdemes lenne részletesen és rendszeresen elemezni ezeket. Ha rendelkezésünkre állnának az eredeti adatok és az adott városokban a fertőzöttek, a kórházban ápoltak száma napi bontásban, akkor az itt illusztrációként bemutatottnál összehasonlíthatatlanul pontosabb és részletesebb elemzésre nyílna mód. Olyan elemzésre, amely alapján lehetőség adódna területenként, akár városonként eltérő járványügyi intézkedésekre (egy-egy területen, ahogy látszólag nem nő és alacsony az örökítő anyag koncentrációja, lehetőség lenne a teljes „lockdown” és a gazdasági aktivitás korlátozásának elkerülésére). Adatokkal, elemzésekkel lehet mérsékelni a korlátozásokkal elkerülhetetlenül együtt járó negatív gazdasági hatásokat, a munkanélküliség növekedését, a vállalati teljesítmény és a lakossági fogyasztás csökkenését.

A 2. ábrán a *J* mutató időleges csökkenése, amelyre szeptember 28-a és október 11-e között került sor, a járvány terjedésének mérséklődésére utal. Ezt követően, október 18. és november 1. között, a *J* mutató újra meredek növekedésbe fordul. Ezek szerint a kormánynak már október 19-én, aztán október 26-án újra látnia kellett, hogy a vírus örökítő anyagának koncentrációja *országosan* az időleges csökkenés után erőteljes növekedésnek indult – tehát lépnie kell. Erre a lépésre két héttel később került sor.

2. ábra: A COVID-19 örökítőanyagának becsült szintje a szennyvízben (*J* mutató, 30. hét = 100%), 2020 július 20. - november 15.



3. ábra: A COVID-19 örökítő anyaga a szennyvízben (J mutató, 30. hét = 100%, eredeti idősor és t+6 héttel ábrázolva) és a járvány halálos áldozatainak száma hetente (fő), 2020. július 20. - december 27.

