

13Settembre

## La condivisione dei dati scientifici attraverso i "preprint"

*Dando alle persone il potere della condivisione,  
stiamo rendendo il mondo più trasparente*  
Mark Zuckerberg

Una cosa è chiara. Secondo il database bibliometrico *Dimensions* la pandemia ha provocato una valanga di nuovi articoli: più di **530.000**, pubblicati su riviste o come *preprint*.

Questo ha determinato in un anno il più grande incremento di tutti gli articoli accademici e in assoluto il più grande di sempre.

A determinare questa "esplosione" è stata un interesse improvviso per i *preprint* impressa da "ricercatori medici". Vari server hanno accettato ed accettano documenti direttamente dagli autori, rendendoli disponibili quasi contemporaneamente dopo alcuni controlli essenziali per smistarli senza un parere "peer review" a giornali di buon livello.

Un server leader, **medRxiv**, a sua volta un nuovo arrivato nella galassia delle riviste scientifiche, fondato solo nel 2019, ha pubblicato circa **200 preprint** su vari argomenti solo nel gennaio 2020; entro maggio 2020, il conteggio mensile era arrivato a circa 2000, circa tre quarti di questi erano relativi alla pandemia. Il traffico verso **medRxiv** è salito alle stelle in tandem, con gli abstract visualizzati fino a 6,1 milioni di volte a luglio, rispetto ai 30.000 di agosto 2019.

Alcuni *preprint* hanno fornito risultati utilizzati immediatamente da alcuni clini al "al capezzale dei pazienti". La notizia che il *corticosteroide desametasone* poteva ridurre i decessi in malattie gravi è stata pubblicata come pre stampa nel giugno 2020, ma non in una rivista peer-reviewed fino a febbraio. A quel punto, i medici avevano già somministrato lo steroide per mesi.

Una preoccupazione ricorrente sui *preprint* nella pratica medica era nella mancanza di controllo, ma la pandemia ha chiarito che nemmeno la revisione tra pari era infallibile. Articoli di riviste di alto profilo e sottoposti a revisione paritaria su COVID-19 sono stati ritirati; sono sorti sospetti sul fatto che siano stati fabbricati dati per un lavoro che affermava che l'*idrossiclorochina*, un farmaco antimalarico promosso dall'ex presidente **Donald Trump**, non solo **era inefficace**, come hanno dimostrato molti studi, ma potrebbe danneggiare i pazienti COVID-19. Il confine tra la conoscenza scientifica convalidata, il terreno tradizionalmente rivendicato dalle riviste peer-reviewed, e la conoscenza scientifica non ancora convalidata sotto forma di *preprint* è diventato più sfocato durante la pandemia.

Il COVID-19 ha anche accelerato gli sforzi per creare una via di mezzo per i preprint, sottoponendone alcuni a una sorta di rapida revisione tra pari. Ma questi sforzi rimangono di piccola scala. **Recensioni rapide: COVID-19**, una rivista lanciata da *MIT Press* nell'agosto 2020, ha pubblicato rapidamente recensioni online di circa 150 *preprint* su COVID-19

Ma le previsioni iniziali, secondo cui i *preprint* avrebbero dominato la letteratura sul COVID-19 e avrebbero influenzato tutta la pubblicistica scientifica dopo la pandemia, finora sembrano errate. Fino a maggio, secondo una stima di **Ludo Waltman e Stephen Pinfield** del **Research on Research Institute**, sulla base dei database bibliometrici *Dimensions* e *Unpaywall*, solo il 5% circa di tutti gli articoli di riviste sottoposti a revisione paritaria sugli articoli COVID-19 è iniziato come

prestampa. . (La maggior parte dei *preprint* COVID-19 non è apparsa su riviste peer-reviewed, almeno finora.) E l'impennata di *preprint* dell'anno scorso sta diminuendo, con il totale mensile di medRxiv ora a circa *600 articoli*, circa la metà su COVID-19 .

"Molti ricercatori non sono abituati alle pre stampa e non vengono ricompensati per loro", afferma Waltman. *"Forse i finanziatori avrebbero potuto davvero incoraggiare o addirittura imporre la pre stampa. Se vogliamo davvero [più preprint], le diverse parti interessate devono lavorare insieme. ... Altrimenti, rimarrà un fenomeno su piccola scala"*.

A complemento dello slancio impresso dai preprint, gli editori delle riviste tradizionali hanno promesso di accelerare le revisioni dei manoscritti COVID-19. Hanno avuto un certo successo: nel gennaio 2020, i manoscritti su COVID-19 e argomenti correlati sono stati accettati in media circa 130 giorni dopo l'invio. A luglio 2020, il ritardo era sceso a circa 90 giorni secondo un'analisi di **Daniel Hook** e colleghi di *Digital Science*, basata sul suo database Dimensions. (Lo studio ha esaminato il 20% di tutti i documenti in cui sono state fornite queste date.)

Questo ha accelerato la qualità smussata? Sebbene manchino prove complete, solo lo **0,03%** circa degli articoli di riviste su COVID-19 è stato ritirato, leggermente meno che in tutte le discipline prima della pandemia.

Per quanto riguarda la richiesta di **Wellcome** di rendere gli articoli e i dati COVID-19 gratuiti per tutti, i risultati sono contrastanti. All'inizio del 2020, molti editori si sono impegnati a esentare i documenti COVID-19 dai loro paywall e, dalla scorsa settimana, il **77%** di tutti questi articoli erano liberi di leggere, secondo il database Dimensions. Si tratta di un calo da un massimo dell'**85%** nel maggio 2020, ma ancora ben al di sopra del 50% per la letteratura scientifica appena pubblicata prima della pandemia.

Le prime analisi suggeriscono meno progressi verso la fornitura di "dati aperti", la condivisione di set di dati con altri scienziati per accelerare il riutilizzo dei dati e aumentare la trasparenza della ricerca. Tra più di 7000 *preprint* di medRxiv su COVID-19 pubblicati durante il 2020, gli autori di solo la metà degli autori si sono offerti di condividere prontamente i dati sottostanti .  
Poter condividere è poesia nella prosa della vita! (Sigmund Freud)

#### **A chi legge**

alcuni dati sono ricavati dal report di **Jeffrey Barinard** approu su *Science* il 4 settembre

