

Stefano Canali

## COVID-19: dati e qualità

In questo articolo discuto l'utilizzo di dati come base delle politiche pubbliche nell'ambito della pandemia da COVID-19. Inizio introducendo il dibattito che si è creato attorno a questo tema, in particolare rispetto a una serie di critiche rivolte alla qualità apparentemente bassa dei dati usati come giustificazione di varie misure di contenimento. Procedo poi mostrando come queste critiche sottendano un approccio 'assolutistico' al concetto di qualità e, sulla base di vari lavori di filosofia della scienza e medicina, sostengo invece che si dovrebbe avere un approccio 'contestuale'. Concludo mostrando come questo approccio si applica al contesto della pandemia.

### 1. Introduzione

La pandemia da COVID-19 è un evento con innumerevoli sfaccettature e aspetti che vanno dal politico all'economico, dal sociale al biologico. Una serie di dibattiti si è perciò creata su pressoché tutti questi aspetti. L'ambito biomedico non è certo stato da meno, se non altro per l'apporto fondamentale che la ricerca biomedica può e deve apportare ad un evento – la pandemia – che è (anche) biomedico. In questo senso sono nate discussioni su come l'impatto delle disuguaglianze sulla salute sia stato esacerbato e aggravato dalla pandemia, sui limiti e i confini da delineare tra sanità pubblica e privata, sugli approcci e paradigmi di cura da incentrare su pazienti o comunità di pazienti, etc.

In questo articolo mi voglio concentrare su un dibattito che ha coinvolto tanto la ricerca biomedica quanto la filosofia della scienza e della medicina, la politica e la riflessione sulle politiche pubbliche. Il dibattito riguarda il tipo e la qualità dei dati da usare come base e giustificazione degli interventi di sanità pubblica nell'ambito della pandemia attorno a domande come le seguenti: Quali dati dovremmo usare per decidere se applicare un lockdown? I dati sui benefici dell'utilizzo delle mascherine sono di qualità sufficiente? In che senso i dati sull'incidenza settimanale di positività possono determinare la chiusura di alcune attività commerciali? Nell'articolo intendo dare una prospettiva con cui approcciare queste domande, sostenendo che i dati scientifici siano artefatti dalla natura contestuale e che dovremmo quindi avere un approccio contestuale rispetto alla determinazione della qualità.

Inizio l'articolo introducendo le critiche alla qualità dei dati su COVID-19, che sono state presentate durante lo sviluppo della pandemia e hanno caratterizzato il ruolo dei dati nella gestione politica della pandemia come un fiasco (sezione 2). Nonostante il merito di aver posto l'attenzione sul tema dei dati e della qualità rispetto a COVID-19, nella sezione 3 mostro come queste critiche abbiano una serie di problemi e usino principi di qualità di difficile applicazione nel caso della pandemia. A questo punto, presento un ragionamento più ampio e sostengo che questo dibattito esemplifichi un modo generale di intendere la qualità – cioè l'idea che la qualità sia una proprietà universale e assoluta dei dati. Tuttavia, nella sezione 4 presento una serie di ragioni sulla base delle quali dovremmo invece avere un approccio contestuale alla qualità dei dati, secondo il quale la qualità dei dati va determinata a partire dal contesto in cui questi dati vanno utilizzati e non indipendentemente da ogni contesto. Ritorno alla pandemia per mostrare cosa voglia dire, in termini più concreti, porre attenzione al contesto e come questo si traduca in indicazioni sull'uso dei dati nel caso di COVID-19.

### 2. Gli interventi su COVID-19 sono basati su dati inaffidabili?

L'idea di produrre dati di qualità come risultato della ricerca scientifica è ovviamente una delle basi della scienza contemporanea, in particolare in connessione all'attività sperimentale e al ruolo sempre più importante dei grandi volumi di dati che si possono accumulare, analizzare e

condividere.<sup>1</sup> Avere la qualità come uno degli obiettivi delle attività di ricerca che coinvolgono i dati è forse ancora più importante nel momento in cui la ricerca è chiamata a fornire delle evidenze su cui basare le politiche pubbliche. La pandemia da COVID-19 è un caso esemplificativo di queste dinamiche, nella misura in cui abbiamo assistito a un ruolo sempre più importante delle scienze biomediche nell'attività di consiglio e direzione degli interventi di contenimento.<sup>2</sup> Richiedere che la ricerca e i dati che produce siano di qualità, soprattutto nel momento in cui questi dati sono usati per indicare delle azioni concrete da applicare, sembra un'ovvietà. Tuttavia, come vedremo, i rapporti tra dati, qualità e interventi di politica pubblica sono tutt'altro che semplici e lineari e sono stati oggetto di un esteso dibattito nel caso della pandemia da COVID-19.

Il dibattito in questione ha avuto inizio pressoché all'inizio dell'espansione della pandemia. Nel marzo del 2020 l'epidemiologo e meta-scienziato John P. Ioannidis pubblicò una serie di contributi molto critici rispetto alle attività di ricerca e interventi di sanità pubblica su COVID-19.<sup>3</sup> Ioannidis arriva a caratterizzare queste attività come un vero e proprio "evidence fiasco", sulla base di una serie di considerazioni sui tipi di dati usati per giustificare le prime misure di lockdown nella primavera del 2020. Un primo punto riguarda i dati sul tasso di letalità di SARS-Cov-2, una misura epidemiologica per conteggiare le morti dovute ad una specifica condizione rispetto alla popolazione di individui affetti da questa condizione. Ad esempio, il tasso di letalità di COVID-19 per la popolazione europea si ottiene dividendo il numero di morti dovute a COVID-19 per il numero di tutte le persone che hanno contratto virus e malattia. Il tasso di letalità è un dato estremamente importante per lo studio di una malattia e per la sanità pubblica, nella misura in cui è usato come base per comprendere l'indice di rischio, ed è particolarmente importante per ottenere un'immagine dinamica dello sviluppo di una pandemia. Secondo Ioannidis, tuttavia, i dati sul tasso di letalità di SARS-Cov-2 sono di bassa qualità e si dovrebbe invece studiare la prevalenza dell'infezione su un campione casuale e ripetere questa operazione a intervalli regolari. Un punto simile riguarda più in generale i dati usati per giustificare il distanziamento sociale.<sup>4</sup> Qui il problema, secondo Ioannidis, è che i dati a favore di queste misure sono inaffidabili perché basati su studi osservazionali, conoscenze qualitative e opinioni esperte, invece che su studi controllati e randomizzati (*randomised-controlled trials*). Al contrario, le misure con maggiore supporto di dati riguardano igiene di base, sanificazione e mascherine.

Questo tipo di critiche, presentate da Ioannidis e altri in maniera simile, hanno avuto un'eco importante a livello accademico, politico e più in generale mediatico. In questo senso, una prima cosa importante da dire è che queste critiche si basano su uno sfondo teorico e metodologico ben preciso. Come mostrato dal filosofo della medicina Jonathan Fuller, si tratta di una posizione tipica della cosiddetta 'medicina basata su prove di efficacia' (più nota con il termine inglese *Evidence-Based Medicine*, EBM), un movimento secondo cui le decisioni in ambito sanitario dovrebbero eliminare le preferenze soggettive e basarsi invece su linee guida standardizzate che indicano la qualità delle evidenze.<sup>5</sup> Più concretamente, negli ultimi vent'anni questo movimento ha portato alla creazione di gerarchie e piramidi, che strutturano vari tipi di evidenza in base alla loro affidabilità e qualità: l'idea è che interventi in ambito clinico e sanitario siano di maggiore efficacia se supportati da evidenze che occupano una posizione più alta nella piramide. E a stare più in alto in queste piramidi sono proprio le evidenze ottenute con processi controllati, randomizzati e sistematici. Si

---

<sup>1</sup> S. Leonelli, *La ricerca scientifica nell'era dei Big Data: Cinque modi in cui i Big Data danneggiano la scienza*, e come salvarla, Maltemi, 2019.

<sup>2</sup> S. Lohse. & K. Bschrir. *The COVID-19 pandemic: a case for epistemic pluralism in public health policy*, in "History and Philosophy of the Life Sciences", 43 (58), 2021.

<sup>3</sup> J.P.A Ioannidis. *A fiasco in the making? As the coronavirus pandemic takes hold, we are making decisions without reliable data*, In "STAT", 2020.

<sup>4</sup> J.P.A Ioannidis. *Coronavirus disease 2019: The harms of exaggerated information and non-evidence-based measures*. In "European Journal of Clinical Investigation", 50, e13222, 2020.

<sup>5</sup> J. Fuller. *Models v. Evidence*. In "Boston Review", 2020. <https://bostonreview.net/science-nature/jonathan-fuller-models-v-evidence>

comprende quindi perché molte misure di contenimento della pandemia siano state criticate per la scarsa qualità: perché non sono basate su evidenze ottenute con studi randomizzati e controllati.

### 3. Critiche inaffidabili: perché l'idea del fiasco non funziona

I dati su COVID-19 sono effettivamente un *evidence fiasco*? Il modo di intendere la qualità come conseguenza di randomizzazione, controllo e sistematicità è legittimo ed estremamente influente nella medicina e nelle politiche pubbliche contemporanee. Tuttavia, questo approccio tipico della EBM è stato analizzato e fortemente criticato in una vasta letteratura filosofica, biomedica, sociologica e storica. Il problema del dibattito sull'*evidence fiasco* è che i punti critici della EBM, evidenziati in questa letteratura, restano validi e sono anzi particolarmente evidenti nell'ambito della pandemia.

Un primo problema riguarda la traduzione pratica di questi principi: nel mondo reale è molto complicato ottenere una randomizzazione perfetta e in molti casi perfezionare la randomizzazione significa diminuire l'applicabilità e efficacia degli studi.<sup>6</sup> Ad esempio, nel caso di COVID-19 sono disponibili relativamente pochi studi clinici randomizzati e controllati e anche questi si occupano principalmente di terapia invece che di contenimento e hanno dimensioni limitate, per cui è complicato rilevare l'efficacia delle misure testate.<sup>7</sup> Sostenere quindi che gli interventi di contenimento dovrebbero essere basati su dati randomizzati e controllati risulta quindi problematico, perché questo tipo di studi esiste in scarsa quantità e non ci fornisce informazioni sul contenimento della pandemia.

In secondo luogo, anche se ci fossero degli studi clinici randomizzati e controllati su misure di contenimento di COVID-19, un altro problema identificato nella letteratura sulla EBM rimane: l'estrapolazione dei dati. Il problema è che i processi di randomizzazione e controllo accrescono la validità dei risultati rispetto a influenze esterne, ma rendono difficile estrapolare e applicare i risultati in contesti diversi da quelli controllati nella fase sperimentale. Questo problema è particolarmente grave nel caso di una pandemia, in cui siamo proprio interessati a capire come misure adottate in certe parti del mondo (il lockdown di Wuhan in Cina) possano essere efficaci e applicabili in altri luoghi (l'Europa).<sup>8</sup> Test randomizzati e controllati possono, in altre parole, fornire evidenze rispetto al funzionamento di un intervento, ma ci dicono poco sul perché e come questo intervento abbia funzionato – e spesso è proprio questo che vogliamo sapere per interventi di salute pubblica.

Un ultimo punto che voglio menzionare qui è più ampio e riguarda l'idea che basare le politiche pubbliche sui dati possa essere neutrale, obiettivo e privo dell'influenza di certi valori. Nella filosofia della scienza contemporanea l'idea di obiettività come sinonimo di neutralità e fattualità è stata per lo più superata e in vari lavori si è mostrato che una serie di valori influenza l'uso dei dati, le inferenze e più in generale la creazione della conoscenza scientifica.<sup>9</sup> Come è evidente in particolare nell'ambito della ricerca biomedica, questi valori possono avere un carattere più o meno epistemico, arrivando a includere per esempio principi di semplicità e chiarezza, ma anche giustizia e pluralità.<sup>10</sup> La presenza di valori come questi non significa che la conoscenza scientifica non possa essere obiettiva o di qualità. Al contrario, la presenza di tanti e diversi valori può rafforzare la ricerca e rendere la conoscenza più obiettiva in quanto risultato sociale di scambi,

<sup>6</sup> J. Worrall. *What evidence in evidence-based medicine?*. In "Philosophy of science", 69(S3), S316-S330, 2002.

<sup>7</sup> H. Bauchner, P.B. Fontanarosa. *Randomized Clinical Trials and COVID-19: Managing Expectations*. In "JAMA", 9;323(22), pp.2262-2263, 2020.

<sup>8</sup> A. Broadbent, B. Smart. *Why a one-size-fits-all approach to COVID-19 could have lethal consequences*, In "The Conversation". <https://theconversation.com/why-a-one-size-fits-all-approach-to-covid-19-could-have-lethal-consequences-134252>

<sup>9</sup> H. Douglas. *Science, policy, and the value-free ideal*. University of Pittsburgh Press, 2009; S. Jukola. *On ideals of objectivity, judgments, and bias in medical research—A comment on Stegenga*, "Studies in History and Philosophy of Biological and Biomedical Sciences", 62, pp. 35– 41, 2017.

<sup>10</sup> J. Biddle, R. Kukla. *The geography of epistemic risk*. In "Exploring inductive risk", ed. K. Elliot & T. Richards. Oxford University Press, 215–237, 2017.

critiche e contesti differenti che contengono una pluralità di punti di vista.<sup>11</sup> Tuttavia, sostenere che le politiche di sanità pubblica si possano limitare a seguire la scienza e usare i dati senza l'influsso di una serie di valori è fuorviante. L'utilizzo dei dati implica una serie di valori, che possono ad esempio mettere l'accento su certe caratteristiche dei dati, giudicare la qualità in modi specifici e portare a scegliere quali dati usare e quali escludere.<sup>12</sup> In quest'ottica intendere la qualità come conseguenza di randomizzazione, controllo e sistematicità è un approccio specifico, non è neutrale e si basa su specifiche affermazioni di valore. A maggior ragione gli interventi di politica pubblica restano appunto politici: richiedere che questi interventi siano basati su studi clinici randomizzati e controllati è legittimo (al di là delle complicazioni che abbiamo visto), ma di per sé non è un modo più obiettivo o scevro di considerazioni valoriali.<sup>13</sup>

#### 4. Qualità e dati: il contesto conta

Visti i problemi che incontra la visione della EBM nell'applicazione a COVID-19, come dovremmo pensare al rapporto tra dati, qualità e interventi di salute pubblica? In questa sezione voglio concentrarmi su un'idea più generale che sottende il modo in cui intendiamo la qualità, basato su una visione universale e indipendente di qualità, che ritengo problematica. Introduco perciò un approccio diverso, che sottolinea il ruolo del contesto quando valutiamo la qualità.

Uno dei problemi della EBM rispetto a COVID-19 è l'estensione di un punto già criticato nella letteratura, cioè l'idea che un approccio specifico alla qualità possa essere applicato ad altri contesti, indipendentemente dalle loro caratteristiche specifiche. In questo senso l'associazione tra qualità, randomizzazione e controllo, che nel caso della EBM è nato ed è applicato in ambito clinico e farmacologico (con dei limiti), sarebbe ulteriormente applicabile alla sanità pubblica. Al di là dei problemi di questa applicazione, discussi nella sezione precedente, questo approccio sottende l'idea che la qualità sia una proprietà intrinseca e universale dei dati – un dato è di qualità in generale, indipendentemente da come viene usato e in tutti i contesti. Ad un primo sguardo questa sembra un'idea ragionevole: se un dato è di qualità, perché non dovrebbe essere *sempre* di qualità?

Nonostante questa intuizione, vari studi sulla natura e il ruolo dei dati nella ricerca scientifica e nella formazione della conoscenza indicano una posizione contraria. In primo luogo, la filosofa della scienza Saana Jukola ha mostrato come sia estremamente difficile – e in alcuni casi impossibile – estendere un approccio specifico alla qualità a tutta la ricerca biomedica. Per esempio, nell'ambito della ricerca sulla nutrizione è pressoché impossibile applicare principi di randomizzazione e controllo per ragioni epistemiche, pratiche e anche etiche.<sup>14</sup> Pensare che principi nati in un contesto farmacologico possano applicarsi ad altri ambiti senza considerare gli aspetti specifici del contesto nutrizionale è problematico, nella misura in cui questi aspetti locali potrebbero invece suggerire diversi tipi di dati e diversi standard di qualità.

Un secondo tipo di considerazione riguarda un aspetto fondamentale della qualità, cioè i modi in cui si verifica e valuta. L'analisi del filosofo della scienza Jacob Stegenga sugli strumenti di controllo della qualità dei dati clinici mostra come la qualità cambia nel momento in cui si usano strumenti diversi per valutarla e questi strumenti vengono usati da persone diverse.<sup>15</sup> Da questo si evince che la qualità non è una proprietà intrinseca dei dati, che dipende solo dalle loro caratteristiche, ma in realtà cambia a seconda del contesto e punto di vista da cui viene valutata. La qualità dei dati è quindi tanto associata alle caratteristiche strutturali dei dati quanto il risultato

<sup>11</sup> H.E Longino., *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*, Princeton University Press, Princeton, 1990.

<sup>12</sup> H. Strassheim, P. Kettunen. *When does evidence-based policy turn into policy-based evidence? Configurations, contexts and mechanisms*. In "Evidence & Policy: A Journal of Research, Debate and Practice", 10(2), pp. 259– 277, 2014.

<sup>13</sup> S. Canali, S. Jukola. *Näyttöön perustuvan pandemiapolitiikan haaste (The Challenges of Evidence-Based Pandemic Policy)*. In "Tiede & edistys", 3/2020, pp.212-227, 2020.

<sup>14</sup> S. Jukola. *On the evidentiary standards for nutrition advice*. In "Studies in History and Philosophy of Biological and Biomedical Sciences", 73, pp.1–9, 2019.

<sup>15</sup> J. Stegenga. *Down with the Hierarchies*. In "Topoi", 33, 313–322, 2014.

dell'allineamento di queste caratteristiche con le domande, i metodi, le infrastrutture usate in una comunità di ricerca.<sup>16</sup> I dati sono in questo senso di qualità nel momento in cui possono essere usati per usi specifici, rispondono alle domande di ricerca che si vogliono indagare, possono essere integrati con altri dati e conoscenze presenti – non solo nel momento in cui passano dei test indipendenti di qualità.

Questi punti sono ulteriormente supportati da altre considerazioni legate al tema informatico della qualità come proprietà dell'informazione. In questo ambito, la filosofa dell'informazione Phyllis Illari ha presentato un approccio relazionale, che lega le dimensioni del concetto di qualità a obiettivi e usi specifici.<sup>17</sup> Questo è un modo per superare un problema strutturale della visione universale della qualità, che altrimenti finirebbe per essere 'tra incudine e martello': da un lato si intende la qualità come utile in usi specifici; dall'altro c'è un senso in cui informazioni di qualità dovrebbero essere utili sotto molteplici aspetti.

È sulla base di questi risultati che penso sia necessario superare l'idea di qualità come caratteristica universale dei dati e ritengo sia utile un approccio contestuale.<sup>18</sup> Secondo questo approccio, la qualità è appunto una caratteristica contestuale dei dati, che è legata all'utilizzo specifico dei dati. A seconda dell'uso e delle necessità legate all'uso, alcuni tipi di dati possono essere ritenuti di maggiore qualità rispetto ad altri e questo è solo parzialmente legato alle caratteristiche intrinseche dei dati. Ad esempio, nell'ambito epidemiologico dati ottenuti tramite osservazione sono cruciali per comprendere lo sviluppo di certe malattie nel corso della vita di una popolazione. Altri dati, per esempio da test clinici, non ci forniscono queste informazioni, nonostante siano tradizionalmente ritenuti di maggiore qualità.

Questo punto è strettamente legato a una visione dei dati secondo la quale i dati non sono degli specchi oggettivi del mondo. Come mostrato dalla filosofa della scienza Sabina Leonelli, i dati possono fornire svariate informazioni sul mondo: le caratteristiche intrinseche dei dati delimitano il tipo di informazioni, ma non lo determinano completamente. Le informazioni, infatti, cambiano a seconda degli usi che facciamo dei dati, delle domande che abbiamo in mente, di come le interpretiamo.<sup>19</sup> Per esempio, i dati di monitoraggio su una serie di malattie all'interno di una popolazione possono chiaramente essere usati per studiare lo sviluppo epidemiologico di quelle malattie – ma anche per una ricerca demografica sulla natalità, uno studio storico sui metodi usati nell'epidemiologia contemporanea o persino per un'analisi filosofica sul ruolo dei dati nell'epistemologia biomedica.<sup>20</sup> In questo senso la qualità dei dati cambia a seconda degli usi che ne facciamo: sicuramente le caratteristiche dei dati stessi fanno sì che certi dati non abbiano una qualità sufficiente per certi usi, ma anche gli usi stessi ne possono accentuare o diminuire la qualità.

Avere un approccio contestuale vuol dire dare importanza al ruolo che il contesto svolge nel determinare la qualità dei dati. Questo significa che, nella valutazione e misurazione della qualità di certi dati, non si può pensare di concentrarsi solo sulle caratteristiche specifiche dei dati. Come abbiamo visto, la qualità non si determina solo sulla base di queste caratteristiche. Occorre invece concentrarsi sulle caratteristiche del contesto in cui i dati vengono usati e valutarne la qualità di conseguenza. Questo significa, per esempio, che nel caso di una ricerca sullo sviluppo di malattie in una popolazione, la qualità dei dati ottenuti tramite osservazione e monitoraggio va valutata giudicandone l'allineamento rispetto ai fini della ricerca, i metodi usati per analizzare i dati, etc. In questo senso l'approccio contestuale è anche un modo di contrastare l'idea che si possano avere criteri universali e assoluti per valutare la qualità. Come vedremo, questo è un aspetto cruciale nel

<sup>16</sup> S. Leonelli. *Global Data Quality Assessment and the Situated Nature of "Best" Research Practices in Biology*. In "Data Science Journal", 16, 1–11, 2017.

<sup>17</sup> P. Illari. *IQ: Purpose and Dimensions*. In "The Philosophy of Information Quality", ed. da Floridi, L., Illari, P. Springer: Berlino, 2014; pp. 281–302.

<sup>18</sup> S. Canali. *Towards a Contextual Approach to Data Quality*. In "Data" 5(4), 90 – Special Issue "Data Quality and Data Access for Research", 2020.

<sup>19</sup> S. Leonelli. *Data-Centric Biology: A Philosophical Study*. University of Chicago Press, 2016.

<sup>20</sup> S. Canali. *Making Evidential Claims in Epidemiology: Three Strategies for the Study of the Exposome*. In "Studies in History and Philosophy of Biological and Biomedical Sciences", 82, 101248, 2020.

caso dei dati sulla pandemia, sia rispetto alle critiche sull'apparente *evidence fiasco* di COVID-19, sia per quanto riguarda il modo in cui pensare alla qualità in questo contesto.

### 5. Un approccio contestuale a dati e pandemia

Finora ho presentato le critiche alla qualità dei dati usati per giustificare interventi di sanità pubblica sulla pandemia da COVID-19 e ne ho mostrato i limiti. Ho poi usato questa discussione per sostenere che una visione universale della qualità dei dati sia problematica e vada superato con un approccio contestuale. Quello che voglio fare ora, per concludere l'articolo, è tornare al dibattito su COVID-19, usandolo come caso per mostrare più concretamente cosa significa l'approccio contestuale e applicarlo al ruolo dei dati nella ricerca e nella gestione di una pandemia.

Tenere conto del contesto nell'analisi della qualità dei dati è particolarmente importante nel caso di emergenze come la gestione di una pandemia. Nel momento in cui, nel contesto di un'emergenza, abbiamo necessità di agire sulla base di una serie di evidenze scientifiche, possiamo pensare di applicare criteri universali? Credo di no, e in particolare ritengo ci sia una tensione tra la necessità del contesto politico e sanitario per un'azione veloce e certe linee guida universali. A seconda del contesto, la velocità e tempestività dei dati sono fattori rilevanti e da considerare nel momento in cui si valuta la qualità.<sup>21</sup> In questo senso non è chiaro per quale motivo valori epistemici come randomizzazione e controllo debbano essere ritenuti indici di qualità rispetto a dati che non hanno queste caratteristiche ma forniscono indicazioni tempestive – com'è stato nel caso di dati osservativi sul distanziamento sociale. In determinati contesti, a seconda degli usi che vogliamo fare dei dati, potremmo tenere alcuni valori in maggior conto di altri.

Dare importanza al contesto nella valutazione della qualità ha anche dei risvolti più costruttivi. Anzitutto significa dare importanza al contesto concreto in cui i dati vengono creati, raccolti, interpretati e usati. Questo è un passaggio cruciale, ma spesso dimenticato nella pandemia da COVID-19. Ad esempio, un esercizio importante per la gestione dello sviluppo di una pandemia consiste nel comparare dati sul tasso di riproduzione tra zone diverse, con il fine di capire come il contagio si stia sviluppando e quali misure vadano adottate in regioni diverse. Il problema è che, senza informazioni sul contesto, si tratta di un esercizio molto complicato. Senza informazioni su, per esempio, quali test sono stati usati, da quali regioni provengono, quando sono stati processati, su che fasce della popolazione sono stati fatti, etc. la comparazione è pressoché impossibile o quantomeno poco informativa. Questo è uno dei casi in cui abbiamo bisogno di più metadati, cioè i cosiddetti dati secondari che accompagnano le informazioni principali a cui siamo interessati, come il tasso di riproduzione. In un certo senso questo è uno degli ambiti in cui ci sono stati più problemi durante la pandemia da COVID-19 e in cui c'è ancora molto lavoro da fare. Mentre scrivo queste righe è da poco passato un anno dal momento in cui l'Organizzazione Mondiale della Sanità ha definito COVID-19 una pandemia e ancora ci sono grandi problemi a trovare informazioni chiare su questi fattori contestuali: continuiamo a osservare lo scarso coinvolgimento delle autorità sanitarie locali (spesso le uniche a poter fornire informazioni al riguardo), la difficoltà a comunicare tra realtà regionali e centrali all'interno dei paesi europei, e la carente coordinazione tra strategie di test, tracciamento e isolamento.<sup>22</sup>

Voglio infine sottolineare un ultimo punto indicato dall'approccio contestuale e che penso sia importante, soprattutto rispetto ad una serie di decisioni e scelte in merito all'utilizzo dei dati per studiare e intervenire su COVID-19. Abbiamo visto come nella filosofia della scienza contemporanea il ruolo dei valori sia ritenuto uno dei motori dello sviluppo scientifico e sia in gran parte rifiutata un'idea di obiettività come rimozione di giudizi, valori e punti di vista specifici. In questo senso, ci sono diversi valori che svolgono un ruolo centrale rispetto alla creazione, raccolta e interpretazione di dati sulla pandemia. Lungi dall'essere specchi oggettivi della realtà, i dati rispecchiano sì la realtà ma rispecchiano anche i modi in cui li interpretiamo, le domande a cui

<sup>21</sup> D. Steel. *Accepting an epistemically inferior alternative? A comment on Elliott and McKaughan*. In "Philosophy of Science", 83(4), pp. 606– 612, 2016. D. Jadreškić. *Epistemological and Ethical Aspects of Time in Scientific Research*, 2020.

<sup>22</sup> S. Leonelli. *Data Science in Times of Pan(dem)ic*. In "Harvard Data Science Review", 2021.

cerchiamo di rispondere e i valori da cui partiamo e che vogliamo promuovere. Il lavoro delle filosofe della scienza Elisabetta Lalumera e Cristina Amoretti mostra come una serie di considerazioni valoriali siano fondamentali per i modi in cui è contato il numero di morti da COVID-19 e quindi anche la raccolta dati su questi parametri.<sup>23</sup> Ad esempio, la definizione stessa di morte da COVID-19 fornita dall'Organizzazione Mondiale della Sanità in un certo senso racchiude valori e obiettivi di prevenzione e cura nell'ambito della salute pubblica.

Questa osservazione ha almeno due conseguenze su un piano più pratico. Anzitutto sono necessarie trasparenza comunicativa, discussione pubblica e coordinazione effettiva su questi aspetti contestuali. Se valori decisamente diversi tra loro svolgono un ruolo di simile importanza e non ci sono informazioni al riguardo e la coordinazione, l'integrazione e comparazione tra i dati diventa problematica. Inoltre, mettere l'accento su certi valori rispetto ad altri suggerisca che si debbano fare discussioni aperte e chiare sui valori, soprattutto nel momento in cui hanno un influsso così importante su decisioni politiche. Un secondo elemento che voglio sottolineare è che l'influenza dei valori sui dati in realtà apre alla possibilità di promuovere un utilizzo responsabile ed efficace dei dati.<sup>24</sup> La trasparenza e coordinazione su valori, giudizi e altri aspetti contestuali aumenta le possibilità di integrare dati – nel momento in cui sono note le condizioni e i limiti entro i quali i dati sono stati raccolti e usati risulta più facile capire se si possano incrociare con altri.<sup>25</sup> Allo stesso tempo, discutere il ruolo dei valori nei contesti d'uso dei dati è un modo per interrogarsi su quali valori siano meritevoli di promozione e quali siano insiti nell'uso dei dati, implicitamente o meno. Per esempio, una serie di lavori ha recentemente mostrato come dati raccolti da dispositivi e sensori indossabili, come gli *smartwatch*, possano essere usati per predire lo sviluppo di COVID-19 ancor prima dell'insorgenza di sintomi.<sup>26</sup> Si tratta di dati ritenuti di grande qualità, in quanto molto specifici e vicini agli individui, ma anche di dati estremamente personali e privati. Riflettere sul tipo di valori insiti nell'utilizzo di queste tecnologie nel contesto di una pandemia suscita una serie di domande: L'uso di questi dati è un modo eccessivo di monitorare e sorvegliare una popolazione, che lede valori individuali di autonomia e riservatezza? O forse è un modo di privilegiare valori di cura e protezione collettive? Come possiamo far coesistere valori così apparentemente diversi? L'approccio contestuale alla qualità e ai dati, di cui ho cercato di mostrare i meriti in questo articolo, non fornisce risposte chiare a queste domande. Penso tuttavia che sia un punto di partenza importante iniziare a porle, specificandole come quesiti indissolubilmente legati ai dati e includendole in un contesto di riflessioni più ampio sulla salute come valore da pensare e rifondare alla luce della pandemia da COVID-19.

---

<sup>23</sup> M.C. Amoretti, E. Lalumera. *Covid-19 come causa di morte: una nozione tra fatti e valori*. In "Questione e giustizia", 2021. <https://www.questionegiustizia.it/articolo/covid-19-come-causa-di-morte-una-nozione-tra-fatti-e-valori>

<sup>24</sup> L. Bezuidenhout, E. Ratti. *What Does It Mean to Embed Ethics in Data Science? An Integrative Approach Based on Microethics and Virtues*. In "AI & SOCIETY". 2020.

<sup>25</sup> S. Leonelli, N. Tempini. *Where health and environment meet: the use of invariant parameters in big data analysis*. In "Synthese", 2018.

<sup>26</sup> T. Mishra, M. Wang, A.A. Metwally, et al. *Pre-symptomatic detection of COVID-19 from smartwatch data*. In "Nature Biomedical Engineering", 4(12), pp. 1208-1220, 2020.

## Bibliografia

- Amoretti, M.C., Lalumera, E. *Covid-19 come causa di morte: una nozione tra fatti e valori*. In "Questione e giustizia", 2021. <https://www.questionegiustizia.it/articolo/covid-19-come-causa-di-morte-una-nozione-tra-fatti-e-valori>
- Bauchner H, Fontanarosa P.B. *Randomized Clinical Trials and COVID-19: Managing Expectations*. In "JAMA", 9;323(22), pp.2262-2263, 2020.
- Bezuidenhout, L. & Ratti, R. *What Does It Mean to Embed Ethics in Data Science? An Integrative Approach Based on Microethics and Virtues*. In "AI & SOCIETY", 2020.
- Biddle, J., & Kukla, R. *The geography of epistemic risk*. In "Exploring inductive risk", ed. K. Elliot & T. Richards. Oxford University Press, 215–237, 2017.
- Broadbent, A. & Smart, B. *Why a one-size-fits-all approach to COVID-19 could have lethal consequences*. In "The Conversation". <https://theconversation.com/why-a-one-size-fits-all-approach-to-covid-19-could-have-lethal-consequences-134252>
- Canali, S. *Towards a Contextual Approach to Data Quality*. In "Data" 5(4), 90 – Special Issue "Data Quality and Data Access for Research", 2020.
- Canali, S. *Making Evidential Claims in Epidemiology: Three Strategies for the Study of the Exposome*. In "Studies in History and Philosophy of Biological and Biomedical Sciences", 82, 101248, 2020.
- Canali, S. & Jukola, S. *Näyttöön perustuvan pandemiapolitiikan haaste (The Challenges of Evidence-Based Pandemic Policy)*. In "Tiede & edistys", 3/2020, pp.212-227, 2020.
- Douglas, H. *Science, policy, and the value-free ideal*. University of Pittsburgh Press, 2009
- Fuller, J. *Models v. Evidence*. In "Boston Review", 2020. <https://bostonreview.net/science-nature/jonathan-fuller-models-v-evidence>
- Illari, P. *IQ: Purpose and Dimensions*. In "The Philosophy of Information Quality", ed. da Floridi, L., Illari, P. Springer: Berlino, 2014; pp. 281–302.
- Ioannidis, J.P.A. *A fiasco in the making? As the coronavirus pandemic takes hold, we are making decisions without reliable data*, In "STAT", 2020.
- Ioannidis, J.P.A. *Coronavirus disease 2019: The harms of exaggerated information and non-evidence-based measures*. In "European Journal of Clinical Investigation", 50, e13222, 2020.
- Jadreškić, D. *Epistemological and Ethical Aspects of Time in Scientific Research*, 2020. <https://doi.org/10.15488/10248>
- Jukola, S. *On ideals of objectivity, judgments, and bias in medical research—A comment on Stegenga*, "Studies in History and Philosophy of Biological and Biomedical Sciences", 62, pp. 35– 41, 2017.
- Jukola, S. *On the evidentiary standards for nutrition advice*. In "Studies in History and Philosophy of Biological and Biomedical Sciences", 73, pp.1–9, 2019.
- Leonelli, S. *Data-Centric Biology: A Philosophical Study*. University of Chicago Press, 2016.
- Leonelli, S. *Global Data Quality Assessment and the Situated Nature of "Best" Research Practices in Biology*. In "Data Science Journal", 16, 1–11, 2017.
- Leonelli, S. *La ricerca scientifica nell'era dei Big Data: Cinque modi in cui i Big Data danneggiano la scienza, e come salvarla*, Maltemi, 2019.
- Leonelli, S. *Data Science in Times of Pan(dem)ic*. In "Harvard Data Science Review", 2021.
- Leonelli, S. & Tempini, N. *Where health and environment meet: the use of invariant parameters in big data analysis*. In "Synthese", 2018.
- Lohse, S. & Bschor, K. *The COVID-19 pandemic: a case for epistemic pluralism in public health policy*, in "History and Philosophy of the Life Sciences", 43 (58), 2021.
- Longino, H.E., *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*, Princeton University Press, Princeton, 1990.
- Mishra, T., Wang, M., Metwally, A.A. et al. *Pre-symptomatic detection of COVID-19 from smartwatch data*. In "Nature Biomedical Engineering", 4(12), pp. 1208-1220, 2020.
- Steel, D. *Accepting an epistemically inferior alternative? A comment on Elliott and McKaughan*. In "Philosophy of Science", 83(4), pp. 606– 612, 2016.



Stegenga, J. *Down with the Hierarchies*. In "Topoi", 33, 313–322, 2014.

Strassheim, H. & Kettunen, P. *When does evidence-based policy turn into policy-based evidence? Configurations, contexts and mechanisms*. In "Evidence & Policy: A Journal of Research, Debate and Practice", 10(2), pp. 259– 277, 2014.

Worrall, J. *What evidence in evidence-based medicine?*. In "Philosophy of science", 69(S3), S316-S330, 2002.