مجله‌ی بررسی‌های آمار رسمی ایران
سال ۲۵، شماره‌ی ۲، پاییز و زمستان ۱۳۹۳، صص ۲۱۳–۲۲۹

# نقش داده‌ها در پژوهش و راه‌کارها

باربارا اِی. اندرسون

دانشگاه میشیگان

مترجم: فاطمه تنها[†] و طه نوراللهی[‡,*]

[†] دانشگاه پیام نور، واحد ساوه
[‡] وزارت تعاون، کار و رفاه اجتماعی

**چکیده:** داده‌ها برای پژوهش‌های علمی و برنامه‌ریزی راه‌کارها ضروری هستند. با این حال، لازم است کیفیت داده‌ها و براوردها و مدل‌هایی که بر اساس آن داده‌ها بنا می‌شوند مورد توجه قرار گیرند. علاوه بر این، داده‌ها باید برای آزمایش ایده‌های نو و اعتبارسنجی کار دیگران از راه تکرار، به راحتی در اختیار پژوهش‌گرها قرار گیرند و هم‌زمان باید از پاسخ‌گویانی که داده‌ها را فراهم می‌کنند مراقبت شود. سه مسئله در باره‌ی داده‌ها مورد بحث قرار گرفته است: اول، آماده بودن و درستی داده‌ها برای پژوهش‌های نو و بازتحلیل با مراقبت از انسان‌ها. دوم، دشواری‌های براورد نماگرها بر اساس داده‌های ناقص یا تعمیم‌ناپذیر، و سوم، استفاده از داده‌ها برای ساخت مدل‌هایی برای پیش‌بینی آینده، فرض‌هایی که این مدل‌ها بر آن اساس ساخته می‌شوند و ارزیابی درستی پیش‌بینی‌های گذشته. در هر یک از این زمینه‌ها، توجه بیش‌تر به چگونگی استفاده، تفسیر و ارایه‌ی داده‌ها به جامعه‌ی علمی و سیاست‌گذار ضروری است.

---

Anderson, B.A. (2014). The Role of Data in Research and Policy. Population Studies Center, Report 14-818, Institute for Social Research, University of Michigan.

* نویسنده‌ی عهده‌دار مکاتبات

دریافت: ۱۳۹۴/۲/۲۱، پذیرش: ۱۳۹۴/۶/۲۴.

## مقدمه

داده‌ها برای پژوهش‌های علمی و برنامه‌ریزی راه‌کارها ضروری هستند. با در نظر داشتن این مطلب است که در بسیاری از زمینه‌ها، درخواست برای راه‌کارهای شواهد مبنا افزایش یافته است. با این حال، لازم است به کیفیت داده‌ها و مشخصه‌های براوردها و مدل‌هایی که بر اساس آن داده‌ها شکل گرفته‌اند توجه شود.

سه مورد از دغدغه‌های ویژه عبارت‌اند از:

۱. **آماده بودن و درستی داده‌ها برای پژوهش‌های جدید و بازتحلیل هم‌زمان با مراقبت از** انسان‌ها: به اشتراک گذاشتن داده‌ها برای پژوهش‌های جدید و تکرارپذیر بودن یافته‌ها با مانع‌هایی همراه است و بین دسترسی به داده‌ها و محافظت از انسان‌ها تنش‌های ذاتی وجود دارد.

۲. **دشواری‌های براورد کردن مدل‌ها و نماگرها بر اساس داده‌های ناقص یا تعمیم‌ناپذیر:** نماگرها باید از روی داده‌هایی که به قدر کافی درست و نااریب هستند براورد شوند به‌طوری که تصویری گمراه‌کننده از وضعیت را ارایه ندهند. روش‌های براورد کردن نماگرها اغلب بر اساس رابطه‌ها در زمینه‌هایی تهیه می‌شوند که کیفیت داده‌ها عالی است. این نماگرها ممکن است در زمینه‌هایی که کم‌تر شناخته شده‌اند همان رابطه را با داده‌ها نداشته باشند، اما تشخیص آن بسیار دشوار است. تلاش برای استفاده از داده‌های مناسب‌تر می‌تواند به دشواری‌های بیش‌تر منجر شود.

۳. **استفاده از داده‌ها برای ساخت مدل‌هایی برای پیش‌بینی آینده، فرض‌هایی که این مدل‌ها بر آن اساس ساخته می‌شوند، و ارزیابی درستی پیش‌بینی‌های گذشته:** مدل‌های خط سیر تغییرهای جمعیت‌شناختی به ناچار بر اساس وضعیت‌هایی ساخته می‌شوند که پیش از این تغییر خاصی در آن‌ها رخ داده است. پیش‌بینی آینده به فرض‌هایی نیاز دارد که ممکن است دل‌بخواهی باشند و همچنان که وضعیت در کشورهای پیش‌رفته از لحاظ جمعیت‌شناختی تغییر می‌کند بتوانند به شکلی بنیادی تغییر کنند و وضعیتی ناپایدار را برای کاربران پیش‌بینی‌ها به وجود آورند.

در هر یک از این زمینه‌ها، توجه فزاینده به چگونگی استفاده، تفسیر و ارایه‌ی داده‌ها به جامعه‌ی علمی و سیاست‌گذاری ضروری است.

## پژوهش‌های بنیادی

برخی از اصل‌های اساسی گرداوری و تحلیل داده‌ها عبارت‌اند از:

۱- **داده‌ها باید برای تکرار کار دیگران و آزمون ایده‌های نو آماده باشند:** استانداردی فزاینده وجود دارد که باید داده‌ها برای بازتحلیل در اختیار افراد دیگر در جامعه‌ی علمی قرار گیرند و شماری روزافزون از مجله‌های علمی درخواست می‌کنند که دسترسی به داده‌های مورد استفاده در مقاله‌ها در یک مخزن داده‌ها فراهم شود [۳]. داده‌هایی که با استفاده از بودجه‌ی دولت امریکا گرداوری می‌شوند قرار است معمولاً ظرف مدت دو سال از پایان گرداوری داده‌ها، در اختیار پژوهش‌گرهای دیگر قرار گیرند.

۲- **هویت، پاسخ‌ها و داده‌های دیگر در مورد پاسخ‌گوها باید به‌طور مؤثر محافظت شوند:** نگرانی در باره‌ی مراقبت از انسان‌ها در واکنش به آزمایش‌های پزشکی نازی‌ها و همچنین افشای نقض فاحش حقوق بیماران در امریکا افزایش یافت. در مطالعه‌ی تاسکیگی، از سال ۱۹۳۲ تا ۱۹۷۲ تعداد ۶۰۰ مرد فقیر آمریکایی- آفریقایی در تاسکیگی، آلاباما، مورد بررسی قرار گرفتند که ۳۹۹ نفر از آن‌ها از قبل سیفلیس داشتند. به شرکت‌کنندگان گفته نشد که سیفلیس دارند و برای سیفلیس درمان نشدند حتی پس از این که پنی‌سیلین به‌عنوان درمان مؤثر سیفلیس در سال ۱۹۴۷ پذیرفته شد. بسیاری از شرکت‌کنندگان و همسران و فرزندان آن‌ها از بیماری سیفلیس جان باختند. در سال ۱۹۷۲، اطلاعاتی درباره‌ی این مطالعه به بیرون درز کرد که منجر به خاتمه دادن به این مطالعه شد [۱۴] و [۱۸]. انزجار از مطالعه‌ی تاسکیگی انگیزه‌ای برای گزارش بلمونت شد که مبنایی برای دستورالعمل‌های هیئت بازبینی نهادی است که بر همه‌ی پژوهش‌های علمی در امریکا نظارت دارد.

۳- **تا حدی دانستن بهتر از هیچ ندانستن در یک وضعیت معین است:** مدل‌های رفتاری اغلب بر گرد هم آمدن همه‌ی داده‌های آماده‌ی با کیفیت عالی استوارند. پس مدلی که به داده‌ها برازش می‌شود باید بتواند در وضعیت‌های دیگر نیز مورد استفاده قرار گیرد. اگر این مدل به‌منظور استفاده در براورد نماگرها برای وضعیت‌هایی در نظر گرفته شده باشد که مناسب‌ترین داده‌ها برای آن وضعیت آماده نیست، پس داده‌های ورودی در وضعیتی که کسی بخواهد براوردهایی تولید کند باید به آسانی در دسترس

باشند. همیشه میان گنجاندن داده‌های هر چه بیش‌تر برای ساخت مدل و اطمینان از گنجانده نشدن داده‌های بی‌کیفیت یا اریب تنش وجود دارد.
یک نمونه از آن، مدل جدول‌های عمر منطقه‌ای کول دمنی است [۱۲]. این مدل‌های مرگ و میر بر اساس ۳۲۶ جدول عمر که تصور می‌شد داده‌هایی با کیفیت خوب دارند بنا شد. از این تعداد، ۳۲۴ جدول مربوط به اروپا و امریکای شمالی بود. جدول‌های عمر با مرگ و میر بسیار زیاد تقریباً همه مربوط به اروپای تاریخی بودند. این آگاهی وجود داشت که جدول‌های عمر استفاده‌شده ممکن است همه‌ی تجربه‌ی جهان را نشان ندهند اما گنجاندن جدول‌های عمر مربوط به پرسش از کم درامدها که ممکن بود داده‌هایی بی‌کیفیت داشته باشند بسیار مخاطره‌آمیز به‌نظر می‌رسید.

۴- **گاهی تنها راهنمای آینده آن چیزی است که در گذشته اتفاق افتاده است:** مدل‌بندی مرگ و میر که در موضوع قبلی مورد بحث قرار گرفت دشواری‌های خود را دارد، اما مدل‌بندی باروری می‌تواند مخاطره‌آمیزتر هم باشد. برنامه‌های تنظیم خانواده در تایوان و جاهای دیگر در بسیاری از قسمت‌های آسیا در دهه‌های ۱۹۶۰ و ۱۹۷۰ کاملاً موفق بودند. به‌نظر می‌رسید که این برنامه‌ها نشان داده‌اند که هرگاه وسیله‌های ضد بارداری مؤثر با کاربری آسان در دسترس باشد باروری به سرعت کم می‌شود. این دیدگاه هنگامی زیر سؤال رفت که باروری در کشورهای جنوب صحرای آفریقا و برخی از بخش‌های امریکای لاتین نسبت به تغییر به مراتب مقاوم‌تر به‌نظر رسیدند [۹].

پیش‌بینی‌های جمعیتی به خودی خود برای برنامه‌ریزی آینده مهم‌اند اما بخشی از اطلاعات ورودی برای بسیاری از هدف‌های دیگر مانند براورد کافی بودن عرضه‌ی مواد غذایی در یک منطقه در آینده را نیز تشکیل می‌دهند. نمونه‌هایی شناخته‌شده در گذشته وجود داشته‌اند که به علت آینده‌ی تغییرپذیر ناشناخته، پیش‌بینی‌ها بسیار پرت بوده‌اند.

## پژوهش‌های پیش‌رفته

پژوهش‌ها به کاستی‌ها و ملاحظات در باره‌ی داده‌ها، مدل‌ها و براوردها اشاره کرده‌اند. این گونه مطالعه‌ها بر کار بیش‌تری که باید در این زمینه انجام شود تأکید می‌کنند.

## آماده بودن و درستی داده‌ها برای پژوهش‌های جدید و بازتحلیل هم‌زمان با مراقبت از انسان‌ها

دسترسی به داده‌ها برای پژوهش‌های اصلی و برای تکرارپذیری یافته‌ها موضوعی است که در کشورهای کم‌درامد و پردرامد همچنان مطرح است. در هر دو حالت، مسئله‌ی احتکار اطلاعات توسط پژوهش‌گر یا سازمان و نیز مسئله‌ی رعایت محرمانه بودن پاسخ‌گوها مطرح می‌شود.

در بسیاری از کشورهای کم‌درامد، نبودن داده‌های جمعیت‌شناختی با کیفیت بالا که بر اساس آن بتوان براوردهای جمعیتی را انجام داد و به رابطه‌های مشترک میان متغیرهای اجتماعی و جمعیت‌شناختی پرداخت همچنان مطرح است. برای پرداختن به این مسئله، ۴۹ مکان نظارتی جمعیت‌شناختی در ۲۰ کشور کم‌درامد به خصوص در کشورهای جنوب صحرای آفریقا ایجاد شده است، این مکان‌ها شبکه‌ی INDEPTH نام‌گذاری شده‌اند. در این مکان‌ها تلاشی سخت برای گرداوری اطلاعات در مورد رویدادهای جمعیت‌شناختی در بازه‌های فراوان صورت می‌گیرد.

در مورد دسترسی به داده‌های آماده‌شده در مکان‌های نظارتی جمعیت‌شناختی نگرانی‌هایی وجود دارد. گرداوری داده‌ها نیازمند تلاشی عظیم است. افرادی که گرداوری داده‌ها را هدایت و اجرا می‌کنند اغلب تمایلی ندارند که داده‌ها را به پژوهش‌گرانی بیرون از جمع خود تحویل دهند که به اندازه‌ی آن‌ها در این زمینه تلاش نکرده‌اند. از سوی دیگر، استانداردهای علمی ایجاب می‌کنند که داده‌ها برای بررسی مستقل و اعتبارسنجی آماده باشند. اگر پژوهش‌گران دیگر به داده‌ها دسترسی نداشته باشند، بررسی توضیحات متفاوت امکان‌پذیر نخواهد بود و می‌تواند به زیر سؤال بردن ارزش نتیجه‌های پژوهش بر اساس داده‌ها منجر شود. به علاوه، گرداوری داده‌ها بسیار زمان‌بر است و پژوهش‌گران اغلب نیاز دارند که به محض این که یک دور کامل شد توجه خود را به دور بعدی گرداوری داده‌ها معطوف کنند. به این ترتیب، بیش‌تر داده‌های به دست آمده از این گونه مکان‌ها تنها به شکل محلی و تا حدی محدود تحلیل می‌شوند. این مسئله‌ها به بحثی داغ در باره‌ی شرایطی منجر شده است که داده‌های به دست آمده از مکان‌های نظارتی جمعیت‌شناختی باید با آن شرایط در اختیار جامعه‌ی علمی بزرگ‌تر قرار گیرند ([۴]، [۱۰] و [۱۱]).

تعادل بین مراقبت از پاسخ‌گویان، دادن شانس عادلانه برای سود بردن از انتشار داده‌ها به کسانی که آن داده‌ها را گرداوری کرده‌اند و فراهم کردن امکان تکرارپذیری همچنان منشأ

تنش باقی مانده است. در مکان‌های پژوهشی INDEPTH در بسیاری از کشورهای کم‌درامد، بحثی در باره‌ی این که آیا در اصل، پژوهش‌گرهایی که در گرداوری داده‌ها مشارکت نداشته‌اند باید دسترسی داشته باشند یا نه و این که آیا هدف‌های پژوهشی آن‌ها باید با هدف‌های پژوهش‌گران طرح هم‌سو باشد و آیا لازم است که با پژوهش‌گران طرح همکاری داشته باشند در جریان بوده است. این نیازهای بالقوه با اصول ارزیابی مستقل و تکرارپذیری در علم مغایرت دارد.

در امریکا نگرانی در باره‌ی به اشتراک گذاشتن داده‌ها همچنان باقی است. حتی با این که مجله‌ها بیش از پیش به قرار دادن اطلاعات در یک بایگانی نیاز پیدا می‌کنند. این موضوع دست کم تا سال ۲۰۰۹ بر بسیاری از مجله‌های علمی آزاد اینترنتی که طرفدار پژوهش کاملاً باز هستند تأثیری نداشته است [۲۱]. مک کالو و همکارانش در سال ۲۰۰۸ دریافتند که حتی زمانی که داده‌ها در یک آرشیو قرار داده شده بودند، تنها تکرار ۱۴ مقاله از ۶۲ مقاله امکان‌پذیر بود [۲۲].

همچنان که تحلیل‌ها به‌منظور تعیین راه‌کارها بیش از پیش بر اساس آمارگیری‌ها و ریزداده‌های به دست آمده از سرشماری‌ها انجام می‌شوند، نگرانی‌هایی در مورد محافظت از هویت افراد به وجود آمده است. این موضوع اغلب به درون‌آوری داده‌ها به روش‌های گوناگون منجر می‌شود، از جمله پرشیدگی $n$ داده‌ها به وسیله‌ی معرفی یک عامل تصادفی، گروه‌بندی تاریخ‌های رخداد پیشامدها به دوره‌های ۵ ساله یا ۱۰ ساله و درون‌آوری کل مکان‌های جغرافیایی در یک منطقه‌ی نسبتاً بزرگ. هر چند که در تحلیل داده‌های آماده برای استفاده‌ی عموم، این کارها برای محافظت از پاسخ‌گو می‌تواند به نتیجه‌گیری‌های نادرست منجر شود.

گاهی یک تعدیل جزئی در سن افراد با معرفی کردن یک عامل تصادفی می‌تواند به نتیجه‌های نابهنجار مانند نسبت جنسیتی نامعقول منجر شود ([۱] و [۳۲]). تنظیم داده‌های گزارش‌شده به‌منظور درون‌آوری هویت افرادی که بیش از ۶۵ سال سن دارند می‌تواند به براوردهای نادرست مشخصه‌های سالخوردگان، مانند درامد آن‌ها منجر شود. در مورد تحلیل تاریخچه‌ی پیشامدها، تاریخ‌گذاری جزئیات پیشامدها و ترتیب دنباله‌ای آن‌ها مهم است که با انبوهش زمان‌های رخداد امکان‌پذیر نخواهد بود [۱۶]. یک روش مشترک برای درون‌آوری هویت با گروه‌بندی زمان به دوره‌های ۵ ساله یا ۱۰ ساله و

گزارش کردن داده‌ها برای واحدهای جغرافیایی نسبتاً بزرگ ایجاب می‌کند که در مورد چگونگی درون‌آوری هویت پاسخ‌گو باز اندیشی شود.
درون‌آوری جزئیات جغرافیایی می‌تواند به محافظت از پاسخگویان کمک کند [۲۳]. با این حال، پژوهش‌گران اطلاعات تفصیلی در باره‌ی ویژگی‌های منطقه‌های کوچک جغرافیایی را به‌منظور شناسایی خوشه‌های افراد مبتلا به بیماری‌های خاص یا مطالعه‌ی رفتار یا نگرش‌ها، به شکلی فزاینده ادغام کرده‌اند، چون برای این کار به اطلاعات جغرافیایی با جزئیات بسیار زیاد نیاز دارند [۸]. پژوهش‌گر می‌تواند از هیئتی که داده‌ها را کنترل می‌کند اطلاعات تفصیلی‌تر را درخواست کند. اگر هیئت کنترل‌کننده‌ی داده‌ها پژوهش پیش‌نهادی را به اندازه‌ی کافی ارزشمند ببیند، پژوهش‌گر می‌تواند داده‌های تفصیلی‌تر را به‌دست آورد، اما فرایند پذیرش می‌تواند بسیار طولانی و اغلب ناموفق باشد.

## دشواری‌های براورد کردن نماگرها بر اساس داده‌های ناقص یا تعمیم‌ناپذیر

براورد تعداد افراد مبتلا به یک بیماری گاهی بر اساس نتیجه‌های یک آمارگیری است. برای دقیق بودن براوردها باید پاسخ‌گویان آمارگیری نمایانگر کل جمعیت باشند، یا نحوه‌ی تفاوت پاسخ‌گویان با کل جمعیت باید به خوبی درک شده باشد تا بتوان براوردها را برای کل جمعیت ساخت.
برنامه‌ی ایدز سازمان ملل متحد در براورد خود از تعداد افراد با HIV مثبت در هند تجدید نظری نزولی از ۵/۷ میلیون نفر در سال ۲۰۰۶ به ۲/۵ میلیون نفر برای سال ۲۰۰۷ داشت. علت این تجدید نظر نزولی، کاهش بسیار زیاد HIV نبود، بلکه بیش‌تر ناشی از تغییر در پایه‌ی براوردها بود. برنامه‌ی ایدز سازمان ملل پایه‌ی براوردهای خود را از داده‌های بالینی گروه‌های پرخطر درمانگاه‌ها (زنان باردار، مصرف کنندگان مواد مخدر تزریقی، کارکنان جنسی تجاری) به یک آمارگیری جمعیت مبنای نمایانگر تغییر داد. واضح بود که در براوردهای قبلی، شیوع ایدز در جمعیت عمومی هند به میزان زیادی بیش‌براورد شده بود ([۲۵]، [۳۰] و [۳۱]).
براوردهای جدید به وضوح درست‌تر از براوردهای قدیمی است. با این حال، بعضی‌ها از این تغییر خشنود نیستند، چرا که فکر می‌کنند این براورد جدید ممکن است به توجه و تخصیص پول کم‌تر برای مبارزه با HIV منجر شود. همچنین، برخی این تغییر گزارش‌شده را واقعی تفسیر کردند و بنا بر این در میزان کاهش واقعی ایدز مبالغه شد. این مثال

نشان می‌دهد که چگونگی گرداوری داده‌ها و چگونگی انتخاب پاسخ‌گویان آمارگیری برای گرداوری داده‌ها می‌تواند تأثیر زیادی در نتیجه‌گیری از آن داشته باشد.
سازمان ملل در سال ۱۹۸۲ مدل‌هایی از الگوهای مرگ و میر را بر اساس داده‌هایی از ۲۲ کشور کم‌تر توسعه‌یافته ساخته است [۲۷]. این مدل‌های جدید مرگ و میر برای بهبود مدل‌های کول- دمنی که بر اساس داده‌هایی از اروپا و آمریکای شمالی بود در نظر گرفته شده بودند. این مدل‌های مرگ و میر بیش‌تر برای وضعیت کشورهای کم‌درآمد در نظر گرفته شده بودند. سازمان ملل یک مدل کلی بر اساس داده‌هایی از هر ۲۲ کشور و همچنین ۴ مدل اضافی بر اساس داده‌هایی از زیرمجموعه‌های کشورها ایجاد کرد. متأسفانه، بعدها به این نتیجه رسیدند که مدل مردان برای الگوی امریکای لاتین در اصل مدلی از خطای داده‌ها، ناشی از اشکال‌هایی در داده‌های مربوط به مردان سالخورده‌تر در جدول‌های عمر ارایه‌شده بوده است [۱۳].

## درستی پیش‌بینی‌های گذشته، استفاده از داده‌ها برای ساخت مدل‌هایی به‌منظور پیش‌بینی آینده و فرض‌هایی که این مدل‌ها بر آن اساس ساخته می‌شوند.

با وجود اهمیت پیش‌بینی‌های جمعیتی، کار نسبتاً کمی برای ارزیابی درستی آن‌ها انجام شده است. پژوهش کیلمن در سال ۱۹۹۸ در این خصوص جالب به نظر می رسد [۱۹]. او درستی پیش‌بینی‌های سازمان ملل برای سال‌های ۱۹۵۱ تا ۱۹۸۸ را ارزیابی کرده است. گاهی نادرستی به علت خطا در براورد کردن جمعیت در بار نخست بوده است. پس از انتشار نتیجه‌های سرشماری سال ۱۹۵۳ چین در سال ۱۹۵۴، براورد جمعیت جهان برای سال ۱۹۵۰ افزایش داده شد، به این علت که معلوم شد جمعیت چین بیش از ۱۰۰ میلیون نفر بیش‌تر از آن چیزی بود که پیش از آن تصور می‌شد. در زمان‌های دیگر نیز فرض‌های مربوط به آینده نادرست بوده است. در سراسر جهان، کاهش مرگ و میر سریع‌تر از پیش‌بینی‌ها بوده است. همچنین سرعت کاهش باروری پس از دهه‌ی ۱۹۷۰ بیش‌تر از حد انتظار بود که تا اندازه‌ای ناشی از راه‌کارهای اجرا شده به‌علت هشدار در باره‌ی افزایش زیاد نرخ رشد جمعیت در بسیاری از کشورهای کم‌درآمد در دهه‌های ۱۹۶۰ و ۱۹۷۰ بوده است [۱۹].

همچنین کیلمن در سال ۲۰۰۸ نشان داد که در دوره‌ی ۱۹۵۰ تا ۲۰۰۱، پیش‌بینی‌های جمعیتی در مورد وضعیت جمعیت‌شناختی مورد انتظار برای آینده (کل جمعیت، مرگ و میر، باروری، مهاجرت‌های بین‌المللی) که توسط دفترهای آمار ملی اروپا ارایه می‌شد دقیق‌تر نشده بود [۲۰]. با وجود این که مقدار و کیفیت داده‌هایی که این پیش‌بینی‌ها بر آن اساس پایه‌گذاری شده بودند بهبود یافته بود، این مسئله واقعیت داشت.

بخش جمعیت سازمان ملل، تولیدکننده‌ی اصلی براوردها و پیش‌بینی‌های معتبر کل جمعیت و فرایندهای جمعیت‌شناختی مانند مرگ و میر و باروری است. مهم‌ترین بخش از پیش‌بینی جمعیت، فرض‌هایی درباره‌ی آینده‌ی باروری است. بین سال‌های ۲۰۰۴ و ۲۰۱۲، بخش جمعیت سازمان ملل متحد فرض پایه‌ای خود را در مورد جریان کاهش باروری به سطحی که نتیجه‌ی آن رشد جمعیت صفر و به دنبال آن مسیر باروری است را ۴ بار تغییر داد ([۲] و [۶]). با مرگ و میر کم، رشد جمعیت صفر نیاز به باروری کل (TFR) ۲٫۰۷ دارد. نرخ باروری کل که نتیجه‌ی نهایی آن رشد جمعیت صفر است باروری جایگزینی نیز نامیده می‌شود.

این تغییرها بر اساس مشاهده‌ی تاریخچه‌ی برخی کشورهای کم‌درامد و پردرامد بود. قبل از سال ۲۰۰۴، بخش جمعیت سازمان ملل متحد مدت‌ها بود که پیش‌بینی می‌کردTFR مجانباً به سطح جایگزینی (TFR= ۲٫۰۷) می‌رسد. فرض بر این بود که در نهایت همه‌ی کشورها دارای جمعیت‌های مانا با مرگ و میر کم و باروری کم خواهند بود.

در دهه‌ی ۱۹۹۰، در بسیاری از کشورها، نرخ باروری زیر سطح جایگزینی (TFR<۲٫۰۷) تداوم یافته بود، که گاهی به پایین‌ترین سطح باروری (TFR<-۱٫۳) می‌رسید. پس از مذاکره‌ها و همفکری‌های گسترده، فرض‌های پیش‌بینی باروری در سال ۲۰۰۴ تغییر کرد و پیش‌بینی شد که نرخ باروری کل همه‌ی کشورها مجانباً به ۱٫۸۶ نزدیک خواهد شد که به معنی کاهش بلند مدت جمعیت بود. این یک انحراف بزرگ از فرض قبلی رشد جمعیت نهایی صفر بود.

در دهه‌ی ۲۰۰۰، TFR حداقل در سه دوره‌ی پنج ساله در ۲۱ کشوری که باروری زیر سطح جایگزینی داشتند افزایش یافت. بر اساس افزایش باروی در این کشورها، در سال ۲۰۱۰ فرض‌ها دوباره تغییر یافتند، به گونه‌ای که در مدل جدید، TFR در کشورهای دارای باروری زیر سطح جایگزینی به سمت جایگزینی افزایش یافت به نحوی که هر چه TFR بیش‌تر زیر سطح جایگزینی بود شتاب افزایش سریع‌تر می‌شد. برای کشورهایی با

باروری بالای سطح جایگزنی در سال‌های ۲۰۰۵ تا ۲۰۱۰، مانند الجزایر، پیش‌بینی شده بود که TFR به زیر سطح جایگزینی سقوط می‌کند و سپس به سطح جایگزینی افزایش می‌یابد. این نشانه‌ی بازگشت به فرض جمعیت مانای نهایی بود.
در مورد غیرمنطقی بودن پیش‌بینی‌های باروری برای برخی کشورهای آسیایی که باروری در آن‌ها بسیار کم بود و هیچ نشانه‌ای از افزایش باروری وجود نداشت ([۵]، [۶] و [۷]) فریاد اعتراض بلند بود.
تا حدودی در پاسخ به شکایت‌های مربوط به براوردهای سال ۲۰۱۰ بود که در سال ۲۰۱۲ فرض‌های پیش‌بینی TFR دوباره تغییر کرد. تا سال ۲۰۱۲، TFR دست کم در سه دوره‌ی ۵ ساله در ۲۵ کشور با باروری کم افزایش یافته بود. مدل جدید پیش‌بینی باروی کم برای بسیاری از کشورها بر اساس تجربه‌ی این ۲۵ کشور و سابقه‌ی TFR در هر یک از آن‌ها بود.
پیش‌بینی‌های سال ۲۰۱۲ به انحراف‌هایی بسیار کم‌تر از پیش‌بینی‌های قبلی TFR که بین سال‌های ۲۰۰۴ تا ۲۰۱۰ اتفاق افتاد منجر شده است.

## موضوع‌های کلیدی برای پژوهش‌های آینده

پیش‌رفت موضوع‌هایی که در این جا مورد تأکید قرار گرفتند نیاز به پژوهش‌های علمی دارد ولی نیاز به گفتمان و بحث در فلسفه و اصل‌های اخلاقی نیز وجود دارد. درست مانند هر آزمایشی که در هر جنبه از آن شایستگی‌هایی وجود دارد، ملاحظات و ارزش‌های متضادی هستند که باعث می‌شوند ارایه‌ی یک راه حل کامل در بسیاری از زمینه‌ها تقریبا ممکن نباشد.

## آماده بودن و درستی داده‌ها برای پژوهش‌های جدید و بازتحلیل هم‌زمان با مراقبت از انسان‌ها

علی رغم قائده‌هایی در مورد در دسترس بودن داده‌ها که به‌وسیله‌ی مجله‌ها برقرار شده‌اند، تنوپیر و همکارانش در سال ۲۰۱۱ مستند کرده‌اند که هنوز احتکار داده‌ها رایج است [۲۶]. لازم است بررسی شود که چه گام‌های بیش‌تری می‌توان برداشت تا از عواقب ناخواسته‌ی منفی، مانند دلسردی پژوهش‌گران برای گرداوری داده‌ها نیز جلوگیری شود. به

همین ترتیب، اگر چه شیوه‌ای برای درخواست استفاده‌ی پژوهش‌گران بیرونی از داده‌های INDEPTH وجود دارد، ولی هنوز فرایند آن پیچیده است و زمان می‌خواهد تا معلوم شود که دسترسی چقدر آزاد است.

دفتر سرشماری امریکا در بیش از ۱۵ دانشگاه و مرکز پژوهش، مرکز داده‌های پژوهشی (RDC) ایجاد کرده است. در این مرکزها، پژوهش‌گران دارای پروژه‌های تصویب‌شده می‌توانند در مورد نتیجه‌های اجراهای رایانه‌ای بر مبنای یکایک داده‌های تفصیلی که با این همه جزئیات در مجموعه‌های اطلاعاتی قابل استفاده که برای عموم موجود نیستند تحقیق کنند. این مرکزهای داده‌ها به حل مسئله‌ی دسترسی به داده‌ها و رعایت محرمانه بودن پاسخ‌گویان کمک می‌کنند، اما به‌دلیل این شرط که طرح‌هایی که از مرکز داده‌های پژوهشی استفاده می‌کنند باید «به برنامه‌های دفتر سرشماری سود برسانند» ناقص‌اند [۳۳]. این محدودیت مانعی برای کار آزاد علمی و طیفی از بررسی‌هاست که می‌توانند دنبال شوند. برای این برنامه، تفکر بیش‌تری مورد نیاز است که به نظر می‌رسد تحت تأثیر برخی گرایش‌های مشابهی باشد که مانع به اشتراک گذاشتن داده‌ها میان پژوهش‌گران شده است.

تعادل میان دسترسی به داده‌ها و محافظت از پاسخگویان یک موضوع با ارزش برای راه‌کارهای عمومی است. بحث بیش‌تر میان کسانی که نگران پژوهش‌اند وکسانی که نگران اصل‌های اخلاقی هستند می‌تواند برای روشن کردن آنچه که باید اصل‌های هدایت باشد مثمر ثمر واقع شود. به نظر می‌رسد که این بحث‌ها با تأکید فزاینده بر «داده‌های بزرگ» برای پرداختن به بسیاری از پرسش‌های علمی و راه‌کارها بیش از پیش ضرورت پیدا کرده‌اند ([۲۴] و [۳۴]).

## دشواری‌های براورد کردن نماگرها بر اساس داده‌های ناقص یا تعمیم‌ناپذیر

ساخت مدل‌ها و نماگرها، از جمله برای مرگ و میر که مربوط به وضعیت در بخش‌هایی از جهان است که داده‌هایی برای آن وجود ندارد یا دارای داده‌هایی با کیفیت نامناسب است، همچنان یک چالش باقی مانده است. اشتیاقی قابل درک وجود دارد که همه‌ی داده‌های مناسب با روش‌های رو به پیش‌رفت در ساخت براوردها منظور شوند، ولی خطر گنجاندن داده‌هایی که دارای خطاهای جدی هستند همچنان باقی است.

نیاز به برخی پایه‌ها برای براوردها، از قبیل مرگ و میر، روشن است. یک کشور برای بسیاری از هدف‌های خود به دانستن خیلی چیزها مثل طول عمر متوسط نیاز دارد (که امید به زندگی در بدو تولد نیز نامیده می‌شود). با این حال، در سال ۲۰۱۲، بخش جمعیت سازمان ملل گزارش داد که ۲۶ درصد از کشورهای جهان و ۶۰ درصد از کشورهای آفریقایی داده‌های قابل اعتماد در مورد مرگ و میر بزرگسالان ندارند و دو کشور نیز هیچ‌گونه داده‌های قابل اعتماد در مورد مرگ و میر در هیچ سنی ندارند. در این وضعیت‌ها، براورد بر اساس وضعیت جاهای دیگر ضروری است [۲۹].

**درستی پیش‌بینی‌های گذشته، استفاده از داده‌ها برای ساخت مدل‌هایی به‌منظور پیش‌بینی آینده و فرض‌هایی که این مدل‌ها بر آن اساس ساخته می‌شوند.**

چگونگی استفاده از تجربه‌های گذشته و اندیشه‌ها در باره‌ی آینده یک مشکل اساسی برای مدل‌سازی آینده است. ارزیابی بیش‌تر از درستی پیش‌بینی‌های گذشته برای کل جمعیت و همچنین باروری و مرگ و میر می‌تواند در تصمیم‌گیری‌های آگاهانه‌تر کمک کند. در هر صورت، احتمالاً عاقلانه نخواهد بود که فرض‌ها به روش‌های اساسی بارها تغییر داده شوند، زیرا کاربران نتیجه‌ها به سادگی تصور خواهند کرد که تغییر واقعی روی داده در حالی که هیچ تغییری نبوده است. سازمان ملل در سال ۲۰۰۸ برای سنگاپور پیش‌بینی کرده بود که نرخ باروری کل در سال‌های ۲۰۴۰ تا ۲۰۴۵ به ۱٫۵۹ می‌رسد. در سال ۲۰۱۰، با فرض‌های جدید باروری پیش‌بینی شد که نرخ باروی کل سنگاپور در سال‌های ۲۰۴۰ تا ۲۰۴۵ بیش‌تر از پیش‌بینی تنها دو سال قبل به ۱٫۸۰ می‌رسد. در سال ۲۰۱۲ سازمان ملل پیش‌بینی کرد که باروری کل در سال‌های ۲۰۴۰ تا ۲۰۴۵ به کم‌تر از پیش‌بینی دو سال قبل به ۱٫۳۸ خواهد رسید. در سراسر دوره‌ی ۲۰۰۸ تا ۲۰۱۲ براورد باروری کل در سنگاپور از ۱٫۳۳ به ۱٫۲۵ کاهش یافت. این تغییرها در پیش‌بینی نرخ باروری کل هیچ ربطی به تغییرهای باروری واقعی در سنگاپور نداشت [۲]. آنچه در طول زمان تغییر کرده بود فکر کردن در باره‌ی آینده‌ی باروری، بر اساس خط سیر تغییر در کم‌تر از نیمی از کشورهای دارای باروری کم بود و نه بر اساس تغییری مشاهده‌شده در بسیاری از کشورهایی که باروری آن‌ها پیش‌بینی شده بود.

# مرجع‌ها

[1] Alexander, J.T., Davern, M., and Stevenson, B. (2010). Inaccurate Age and Sex Data in the Census PUMS Files: Evidence and Implications. National Bureau of Economic Research Working Paper No, 15703. Retrieved from http://www.nber.org/papers/w15703.pdf Accessed September 4, 2011.

[2] Anderson, B.A. (2014). Projecting Low Fertility: Some Thoughts about the Plausibility and Implications of Assumptions. University of Michigan Population Studies Center Research Report 14-815 (February) Retrieved from http://www.psc.isr.umich.edu/pubs/pdf/rr14-815.pdf Accessed April 20, 2014.

[3] Anderson, R.G., Greene, W.H., McCullough, B.D. and Vinod, H.D. (2008). The Role of Data/Code Archives in the Future of Economic Research. *Journal of Economic Methodology*, **15**, 99-119.

[4] Baiden, F., Hodgson, A., and Binka, F.N. (2006). Demographic Surveillance Sites and Emerging Challenges in International Health, Editorial. *Bulletin of the World Health Organization*, **86**, 163-164.

[5] Basten, S. (2013a). Re-examining the Fertility Assumptions for Pacific Asia in the UN's 2010 World Population Prospects. University of Oxford Department of Social Policy and Intervention, Barnett Papers in Social Research: 2013/1 (June 7). Retrieved from http://dx.doi.org/10.2139/ssrn.2275938 Accessed June 15, 2013.

[6] Basten, S. (2013b). Comparing Projection Assumptions of Fertility in Six Advanced Asian Economies; or 'Thinking Beyond the Medium Variant.' *Asian Population Studies*, **9**, 322-331.

[7] Basten, S.A., Coleman, D.A., and Gu, B. (2012). Re-examining the Fertility Assumptions in the UN's 2010 World Population Prospects: Intentions and Fertility Recovery in East Asia. a paper presented at the Annual Meeting of the Population Association of America, San

Francisco Retrieved from http://paa2012.princeton.edu/papers/122426 Accessed June 13, 2013.

[8] Berg, M.T., Stewart, E.A., Stewart, E. and Simons, R.L. (2013). A Multilevel Examination of Neighborhood Social Processes and College Enrollment, *Social Problems*, **60**, 513-534

[9] Caldwell, J. and Caldwell, P. (1988). Is the Asian Family Planning Program Model Suited to Africa? *Studies in Family Planning*, **19**, 19-28.

[10] Carrel, M. and Rennie, S. (2008). Demographic and Health Surveillance: Longitudinal Ethical Considerations. *Bulletin of the World Health Organization*, **86**, 612-616.

[11] Chandramohan, D., Shibuya, K., Satel, P., Cairncross, S., Lopez, A.D., Murray, C.D.L., Zaba, B, Snow, R.W. and Binka, F. (2008). Should Data from Demographic Surveillance Systems be Made More Widely Available to researchers? *PLoS Medicine*, **5**, 0169-0170.

[12] Coale, A.J. and Demeny, P. (1966). *Regional Model Life Tables and Stable Populations*. Princeton University Press, Princeton

[13] Dechter, A.R. and Preston, S.H. (1991). Age Misreporting and Its Effect on Adult Mortality Estimates in Latin America. *Population Bulletin of the United Nations*, **31/32**, 1-16.

[14] Fairchild, A.L. and Bayer, R. (1999). Uses and Abuses of Tuskegee. *Science*, **284**, 919-921.

[15] Fisher, T.L. (2010). The Income of the Elderly: The Effect of Changes to Reported Age in the Current Population Survey, a paper presented at the annual conference of the Association for Public Policy Analysis and Management, Boston, October 13 Retrieved from https://www.appam.org/conferences/fall/boston2010/sessions/downloads/4555.1.pdf Accessed July 9, 2011.

[16] Freedman, D., Thornton, A., Camburn, D., Alwin, D. and Young-DeMarco, L. (1988). The Life History Calendar: a Technique for

Collecting Retrospective Event-history Data. *Sociological Methodology*, **18**, 37-68.

[17] INDEPTH Network. (2014). INDEPTH Network: Better Health Data for Better Health Policy. Retrieved from http://www.indepth-network.org/ Accessed April 20, 2014.

[18] Jones, J.H. (1981). *Bad Blood: The Tuskegee Syphilis Experiment*, The Free Press, New York

[19] Keilman, N. (1998). How Accurate Are the United Nations World Population Projections? Population and Development Review, 24 Supplement: Frontiers of Population Forecasting: 15-41.

[20] Keilman, N. (2008). European Demographic Forecasts Have not Become More Accurate over the Past 25 years. *Population and Development Review*, **34**, 137-153.

[21] McCullough, B.D. (2009). Open Access Economics Journals and the Market for Reproducible Economic Research. *Economic Analysis & Policy*, **39**, 117-126.

[22] McCullough, B.D., McGeary, K.A. and Harrison, T.D. (2008). Do Economics Journal Archives Promote Replicable Research? *Canadian Journal of Economics*, **41**, 1406-1420.

[23] Sherman, J.E. and Fetters, T.L. (2007). Confidentiality Concerns with Mapping Survey Data in Reproductive Health Research. *Studies in Family Planning*, **38**, 309-321.

[24] Schuurman, N. (2000). Trouble in the Heartland: GIS and Its Critics in the 1990s. *Progress in Human Geography*, **24**, 569-590.

[25] Steinbrook, R. (2008). HIV in India - A Downsized Epidemic. *New England Journal of Medicine*, **358**, 107-109.

[26] Tenopir, C, Allard, S., Douglas, K., Aydinoglu, A.U. Read, E, Manoff, M. and Frame, M. (2011). Data Sharing by Scientists: Practices and Perceptions, *PloS One*, **6**, 1-21.

[27] United Nations. (1982). *Model Life Tables for Developing Countries*. United Nations, New York.

[28] United Nations. (2011). Assumptions Underlying the 2010 Revision. Retrieved from http://esa.un.org/unpd/World Population Prospects/Documentation/WORLD POPULATION PROSPECTS2010_ASSUMPTIONS_AND_VARIANTS.pdf Accessed July 29, 2011.

[29] United Nations. (2014). World Population Prospects: The 2012 Revision, Methodology of the United Nations population estimates and projections, New York: United Nations Retrieved from http://esa.un.org/unpd/wpp/Documentation/pdf/WPP2012_Methodology.pdf Accessed April 19, 2014.

[30] UNAIDS. (2007). 2.5 Million People Living with HIV in India. Retrieved from http://www.unaids.org/en/KnowledgeCentre/Resources/FeatureStories/archive/2007/20070704_India _new_data.asp Accessed April 19, 2014.

[31] UNAIDS Website. (2008). Q + A on India's Revised AIDS Estimates. Retrieved from http://data.unaids.org/pub/InformationNote/2007/070701_india%20external_qa_en.pdf Accessed June 8, 2010.

[32] U. S. Census Bureau. (2010). Analysis of Perturbed and Unperturbed Age Estimates: 2008. Available at http://www.census.gov/cps/user_note_age_estimates.html Accessed July 2, 2010.

[33] U. S. Census Bureau. (2012). RDC Research Opportunities. Center for Economic Studies Retrieved from https://www.census.gov/ces/rdcresearch/ Accessed April 20, 2014.

[34] United States. White House. Office of Science and Technology Policy. 2012. Obama administration unveils "Big Data" initiative:

Announces $200 million in new R&D investments, Retrieved at http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf Accessed April 19, 2014

## Further Readings

Coale, A.J. and Trussell, T.J. (1996). The Development and Use of Demographic Models, *Population Studies*, **50**, 469-484.

National Academy of Science. (1995). On Being a Scientist: Responsible Conduct of Research, Second Edition, Washington, D. C.: National Academies Press.

Preston, S.H. (1993). The Contours of Demography: Estimates and Projections. *Demography*, **30**, 593-606.

United States, National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research Retrieved from http://www.fda.gov/ohrms/dockets/ac/05/briefing/2005-4178b_09_02_Belmont%20Report.pdf Accessed April 20, 2014.

VanWey, L., Rindfuss, R.R., Gutman, M.P., Entwisle, B. and Balk, D.L. (2005). Confidentiality and Spatially Explicit Data: Concerns and Challenges, Proceedings of the National Academy of Sciences of the United States of America, 102: 15337-15342. Retrieved from http://www.pnas.org/content/102/43/15337.full Accessed April 19, 2014.

**فاطمه تنها**
فوق لیسانس جمعیت‌شناسی
استان مرکزی، ساوه، دانشگاه پیام نور، واحد ساوه.
رایانشانی: f.tanhaa@gmail.com

**طه نورالهی**
فوق لیسانس جمعیت‌شناسی
تهران، خیابان فاطمی، روبروی حجاب، پلاک ۲۵۵، وزارت تعاون، کار و رفاه اجتماعی.
رایانشانی: makannoor@yahoo.com