

EVALUASI PROGRAM KESEHATAN

OLEH:
Grace E. C. Korompis



Penerbit
CV. PATRA MEDIA GRAFINDO
BANDUNG

EVALUASI PROGRAM KESEHATAN

Grace E. C. Korompis



Penerbit
CV. PATRA MEDIA GRAFINDO BANDUNG
2022

EVALUASI PROGRAM KESEHATAN

Penulis : Grace E. C. Korompis

Editing & Layout, desain cover: Tim Patra Media

Hak Cipta @ pada Penulis Dilindungi (All right reserved)

Hak cipta dilindungi undang-undang. Dilarang memperbanyak buku ini sebagian atau seluruhnya, dalam bentuk dan dengan cara apapun juga, baik secara mekanis maupun elektronik, termasuk fotocopy, rekaman dan lain-lain tanpa izin tertulis dari penulis.



Penerbit
CV. PATRA MEDIA GRAFINDO
BANDUNG

Jl. Jend. Sudirman no. 736 - Bandung
Jl. Rorogonggrang Utara II B-10/16 Pharmino
Telp/Fax: 022-6040938 HP: 081214466604
email: patramedia@gmail.com
website: www.patramedia.co.id

Anggota IKAPI

Jenis cetakan : e-book

Tahun publish : November 2022

ISBN 978-623-5481-80-7 (PDF)



KATA PENGANTAR

Syukur kepada Tuhan Yang Maha Kasih atas hikmat dan kebijaksanaan yang melimpah dari-Nya sehingga buku ajar ini dapat diselesaikan dengan baik. Buku ajar ini membahas mengenai Evaluasi Program Kesehatan, dari sisi teori dan hak praktis yang dipraktekkan saat ini.

Penulisan buku ajar ini bermaksud memberikan bahan untuk peserta didik di program studi Ilmu Kesehatan Masyarakat, baik di tingkat sarjana maupun pasca sarjana yang membutuhkan pendalaman mengenai materi perkuliahan yang berkaitan dengan evaluasi kesehatan, sebagai bagian dari manajemen kesehatan secara keseluruhan

Proses penyusunan buku ajar ini melalui beberapa tahap yang diawali dengan melakukan pendalaman teori mengenai evaluasi kesehatan melalui studi literatur, diskusi kelas dalam sesi perkuliahan dan dilanjutkan dengan proses revisi yang melibatkan para pengajar di mata kuliah terkait dan juga mahasiswa.

Untuk itu, penulis mengucapkan terima kasih yang sebesar-besarnya untuk semua pihak yang telah terlibat dalam penulisan buku ajar ini, secara khusus Bidang Akademik Universitas Sam Ratulangi dan Lembaga Pendidikan dan Pengembangan Pengajaran (LP3) Universitas Sam Ratulangi.

Sangat diharapkan bahwa buku ajar ini akan bermanfaat untuk peserta didik dan juga mereka yang membutuhkan pendalaman mengenai proses evaluasi kesehatan. Disampaikan pula permohonan maaf yang sebesar-besarnya apabila masih banyak terdapat kekurangan dalam buku ini. Segala kritik dan saran akan diterima dengan senang hati untuk penyempurnaan buku ini.

Manado, Oktober 2022

Penulis,

dr. Grace E. C. Korompis, MHSM, DrPH

DAFTAR ISI

Kata Pengantar	i
Daftar Isi	ii
PENDAHULUAN	1
Bab 1. Alasan, Proses dan Waktu Pelaksanaan Evaluasi	4
1.1 Alasan Pelaksanaan Evaluasi	4
1.2 Peran Evaluasi	5
1.3 Waktu Evaluasi	7
1.4 Jenis Evaluasi	9
1.5 Kriteria Disain Evaluasi	10
1.6 Pemangku Kepentingan Evaluasi	12
1.7 Tahapan Evaluasi	14
Bab 2. Cakupan Evaluasi	15
2.1 Penetapan Lingkup Evaluasi	15
2.2 Merancang Evaluasi	27
Bab 3. Metode Evaluasi	28
3.1. Pemilihan Metode yang Tepat	28
3.2 Metode Evaluasi Proses dan Dampak	29
3.3 Metode <i>Theory-based Impact Evaluation</i>	30
3.4. Metode <i>Experimental and quasi-experimental Impact Evaluation</i>	32
3.5. Metode <i>Value-for-Money</i>	34
3.6. Metode Sintesis	36
Bab 4. Pengumpulan, Sumber dan Keterkaitan Data	38
4.1. Penentuan Jenis Data	38
4.2 Sumber Data	40
4.3 Kualitas Data	46
4.4. Pengolahan Data	46
Bab 5. Diseminasi Hasil Evaluasi	48
5.1. Pengembangan Rencana Diseminasi	49
5.2. Diseminasi Hasil Evaluasi	49
5.3. Membangun Budaya Evaluasi	50

Bab 6. Evaluasi Program Kesehatan Masyarakat	53
6.1. Alasan Perlunya Mengevaluasi Program Kesehatan Masyarakat	53
6.2. Kerangka Evaluasi Program dalam Kesehatan Masyarakat	54
6.3. Bagaimana Membentuk Tim Evaluasi dan Memilih Lead Evaluator	56
Bab 7. Langkah-langkah Evaluasi Program Kesehatan Masyarakat	58
7.1. Langkah 1: Libatkan Pemangku Kepentingan	58
7.2. Langkah 2: Penjelasan Program	62
7.3. Langkah 3: Disain Evaluasi	66
7.4. Langkah 4: Pengumpulan Data	76
7.5. Langkah 5: Menarik Kesimpulan	78
7.6. Langkah 6: Diseminasi Hasil Evaluasi	80
Glosarium	82
Referensi	86

PENDAHULUAN

Evaluasi memberikan informasi mengenai yang berhasil dan apa yang tidak. Evaluasi harus menjadi penilaian yang ketat dan terstruktur dari keseluruhan kegiatan, intervensi, program atau kebijakan yang telah selesai atau sedang berlangsung. Evaluasi ini penting dalam penentuan besarnya capaian tujuan serta kontribusinya terhadap pengambilan keputusan. Dengan evaluasi, dapat ditentukan apakah suatu intervensi mencapai audiens yang dituju, dilaksanakan sesuai rencana, berdampak sesuai yang diinginkan, apakah hasil yang dicapai lebih baik, juga dapat ditetapkan untuk siapa, bagaimana dan mengapa hal itu berdampak.

Evaluasi melibatkan pengumpulan informasi atau data dan memfasilitasi penilaian tentang keberhasilan dan nilai intervensi. Evaluasi dapat digunakan untuk menginformasikan perubahan untuk meningkatkan intervensi, dan membantu pengambilan keputusan tentang tindakan di masa depan. Evaluasi juga dapat membantu memastikan akuntabilitas publik dan penggunaan sumber daya yang terbatas dapat dilakukan dengan sebaik-baiknya.

Di bidang kesehatan masyarakat, evaluasi dapat bervariasi dalam ukuran dan cakupan. Sebagai contoh, sebuah penelitian mungkin mengevaluasi efek dari kebijakan pemerintah pada ketidaksetaraan kesehatan di suatu negara, atau mengevaluasi apakah penyediaan sandal yang pas mengurangi jatuh di rumah untuk orang tua. Dengan cara ini, evaluasi dapat meningkatkan layanan secara lokal dan memberikan bukti untuk pembuatan kebijakan nasional dan, dengan demikian, meningkatkan praktik kesehatan masyarakat.

Pembahasan selanjutnya dalam buku ini, akan mengupas dengan detail mengenai proses evaluasi yang dilakukan khususnya untuk program kesehatan masyarakat. Uraian mengenai evaluasi dalam buku ini, termasuk yang dilakukan oleh institusi kesehatan formal/ pemerintah di dalam dan luar negeri, organisasi kesehatan baik dalam maupun luar negeri, termasuk oleh organisasi kesehatan dunia/ *World Health Organization* (WHO).

Apa itu evaluasi?

Evaluasi adalah penilaian sistematis dari desain, implementasi dan hasil intervensi. Ini melibatkan pemahaman bagaimana intervensi sedang, atau telah, dilaksanakan dan apa dampaknya, untuk siapa dan mengapa. Ini mengidentifikasi apa yang dapat ditingkatkan dan memperkirakan dampak keseluruhan dan efektivitas biaya.

Kapan evaluasi berguna?

Evaluasi dapat menginformasikan pemikiran sebelum, selama dan setelah implementasi intervensi. Pertanyaan yang berbeda dijawab pada setiap tahap:

- **SEBELUM**

- Apa yang dapat kita pelajari dari evaluasi sebelumnya terhadap intervensi serupa? Bagaimana intervensi diharapkan bekerja? Bagaimana hal itu diharapkan akan disampaikan? Apakah asumsinya valid? Bisakah itu diujicobakan dan diuji sebelum peluncuran penuh? Bisakah peluncuran dirancang untuk memaksimalkan pembelajaran potensial?
- Memberikan bukti yang menginformasikan desain intervensi, cara terbaik untuk mengimplementasikan desain dan kemungkinan hasil yang mungkin terjadi. Membantu mengidentifikasi dan mengurangi ketidakpastian.

- **SELAMA**

- Apakah intervensi berfungsi sebagaimana dimaksud? Apakah itu disampaikan sebagaimana dimaksud? Apa saja dampak yang muncul? Mengapa? Bagaimana itu bisa ditingkatkan? Apakah ada konsekuensi yang tidak diinginkan?
- Memberikan bukti pelaksanaan intervensi dan hasil yang muncul sehingga dapat terus ditingkatkan.

- **SETELAH**

- Apakah intervensi berhasil? Seberapa banyak? Berapa biayanya? Apa yang telah kita pelajari tentang desain dan implementasinya? Apakah perubahan itu berkelanjutan?
- Memberikan bukti tentang desain, implementasi dan hasil, menarik pelajaran untuk masa depan dan memberikan penilaian dampak keseluruhan dari intervensi.

Apa tujuannya?

Ada dua tujuan utama untuk melakukan evaluasi: pembelajaran dan akuntabilitas.

Pembelajaran:

- Untuk membantu mengelola risiko dan ketidakpastian (dari intervensi dan implementasinya);
- Untuk meningkatkan intervensi saat ini dengan memberikan bukti untuk membuat keputusan yang lebih baik
- (dan dimasukkan ke dalam manajemen kinerja dan pekerjaan realisasi manfaat);
- Untuk mendapatkan pemahaman umum tentang apa yang berhasil, untuk

siapa dan kapan, dan menghasilkan contoh untuk pembuatan kebijakan di masa depan;

- Untuk mengembangkan bukti untuk menginformasikan intervensi masa depan.

Akuntabilitas

Penyelenggara program harus akuntabel dan transparan, baik kepada pemerintah dan pemangku kepentingan lainnya. Bukti harus dihasilkan yang dapat menunjukkan dampak intervensi atau hasil yang lebih luas. Bukti efektivitasnya juga diperlukan untuk tinjauan pengeluaran dan sebagai tanggapan terhadap pengawasan dan tantangan dari badan akuntabilitas publik.

Apa arti evaluasi dalam praktik?

Pemantauan dan evaluasi terkait erat, dan evaluasi tipikal akan sangat bergantung pada data pemantauan. Agar dilakukan dengan baik, baik pemantauan maupun evaluasi harus dilakukan selama tahap pengembangan kebijakan dengan keahlian yang terampil untuk memastikan bukti waktu nyata tersedia selama implementasi untuk membantu pengambilan keputusan. Sebuah evaluasi yang komprehensif biasanya akan terdiri dari:

- Analisis dari:
 - o apakah intervensi sedang dilaksanakan sebagaimana dimaksud;
 - o apakah desainnya berfungsi;
 - o apa yang bekerja kurang lebih baik dan mengapa.

Secara bersamaan, jenis pertanyaan ini biasanya disebut sebagai evaluasi proses.

- Tes objektif tentang perubahan apa yang telah terjadi, skala perubahan tersebut, dan penilaian sejauh mana mereka dapat dikaitkan dengan intervensi.

Ini biasanya disebut sebagai evaluasi dampak dan diselidiki melalui pendekatan berbasis teori, eksperimental, dan / atau kuasi-eksperimental.

- Perbandingan manfaat dan biaya intervensi; biasanya disebut sebagai *value-for-money evaluation*/ evaluasi nilai untuk uang.

Untuk memahami sepenuhnya desain, dampak, dan hasil intervensi, semua elemen perlu dieksplorasi.

BAB 1.

ALASAN, PROSES DAN WAKTU PELAKSANAAN EVALUASI

1.1. Alasan Pelaksanaan Evaluasi

Dua alasan utama untuk mengevaluasi adalah pembelajaran dan akuntabilitas.

a. Pembelajaran

Dalam hal pembelajaran, evaluasi dapat memberikan bukti yang dapat digunakan untuk mengelola risiko dan ketidakpastian. Khususnya di bidang-bidang yang inovatif atau terobosan baru, ada kebutuhan akan bukti untuk menggambarkan apakah suatu intervensi berjalan sebagaimana dimaksud. Pembelajaran awal juga dapat menjelaskan bagian mana yang sangat berhasil atau tidak berhasil dan apa yang perlu disesuaikan untuk meningkatkan kinerja. Percontohan dapat berguna dalam konteks ini karena memungkinkan desain, implementasi, dan hasil untuk diuji dalam lingkungan yang terkendali pada skala yang lebih kecil untuk menghasilkan bukti guna menginformasikan inisiatif kebijakan yang lebih luas.

Bahkan area dengan ketidakpastian yang lebih sedikit seringkali dapat mengambil manfaat dari evaluasi: untuk memberikan bukti untuk menginformasikan strategi manajemen manfaat untuk membantu mewujudkan manfaat yang diantisipasi; atau untuk memahami bagaimana memaksimalkan efisiensi dan efektivitas penyampaian. Bahkan ketika kami sangat yakin bahwa intervensi akan efektif, kami setidaknya ingin memantau hasil dan memastikan bahwa mereka sesuai dengan harapan.

Evaluasi juga menghasilkan pembelajaran tentang apa yang berhasil untuk siapa, kapan dan mengapa. Intervensi jarang dilakukan secara terpisah, dan biasanya merupakan salah satu bagian dari program yang lebih besar, membangun apa yang telah dilakukan sebelumnya dan segera diganti dengan ide lain. Penting bagi kita untuk belajar dari intervensi, sehingga kita dapat menerapkan pembelajaran itu pada kebijakan selanjutnya di area yang sama atau area terkait lainnya. Bahkan kebijakan yang dihentikan karena dianggap tidak efektif atau terlalu mahal dapat menghasilkan pembelajaran berharga tentang kesalahan yang harus

dihindari di masa depan, atau mengidentifikasi apakah ada elemen kebijakan yang berhasil.

Pada dasarnya, pembelajaran adalah berkaitan dengan pengambilan keputusan yang baik. Evaluasi dapat memberikan bukti untuk menginformasikan keputusan tentang apakah akan melanjutkan suatu kebijakan, bagaimana meningkatkannya, bagaimana meminimalkan risiko, atau apakah akan berhenti dan berinvestasi di tempat lain.

b. Akuntabilitas

Alasan utama lain untuk mengevaluasi adalah untuk tujuan akuntabilitas. Penyelenggara program (pemerintah) membuat keputusan atas nama rakyat dan membelanjakan pajak yang dikumpulkan dari individu dan bisnis. Pemerintah juga menggunakan inisiatif peraturan, yang berisiko terlalu membebani atau memiliki hasil yang buruk bagi sebagian orang. Pemerintah memiliki tanggung jawab untuk memaksimalkan nilai publik dan hasil yang diberikan untuk uang pembayar pajak dan kegiatan pemerintah. Evaluasi memiliki peran penting dalam hal ini.

Pembuat dan pelaksana program juga harus menginformasikan kepada publik tentang hasil dan nilai dari inisiatif yang mereka lakukan dan bertanggung jawab serta transparan kepada Pejabat Akuntansi untuk pengeluaran mereka. Bukti efektivitas kebijakan juga diperlukan untuk tinjauan pengeluaran dan sebagai tanggapan terhadap pengawasan dan tantangan dari instansi atau badan tertentu.

Evaluasi akuntabilitas cenderung berfokus pada pemantauan dan penilaian dampak. Dalam praktiknya, menyeimbangkan pembelajaran dan akuntabilitas bisa jadi sulit. Akan ada lebih atau kurang fokus pada masing-masing tergantung pada peran yang akan dimainkan evaluasi dengan intervensi khusus yang sedang dipertimbangkan dan kebutuhan pemangku kepentingan.

1.2. Peran Evaluasi

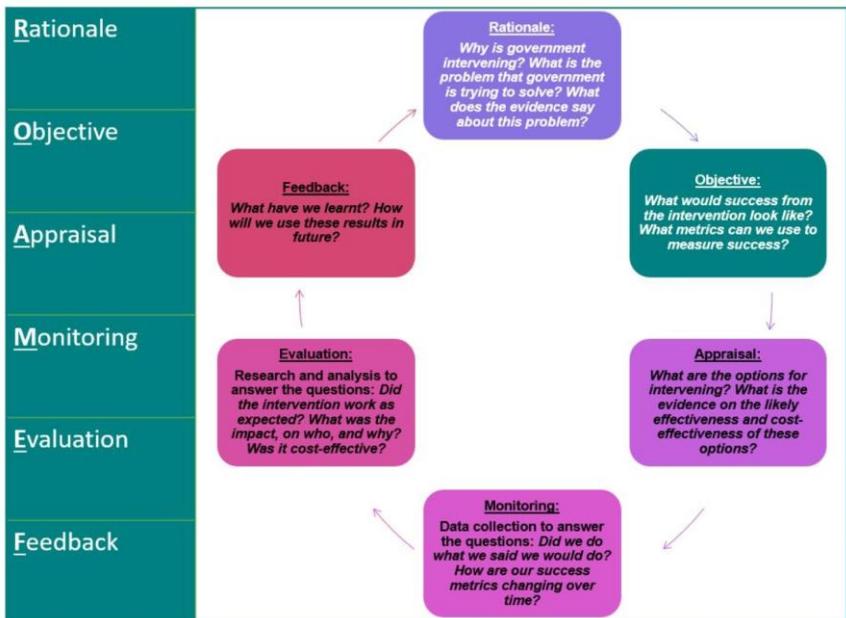
Evaluasi memiliki peran penting yang berkaitan dengan pengambilan kebijakan. Evaluasi berperan dalam semua tahap dalam siklus kebijakan yang dikenal dengan ROAMEF (Gambar 1). Pada praktiknya, ROAMEF adalah cara sederhana untuk

mengekspresikan proses yang kompleks dan evaluasi berguna di semua tahap.

Keluaran dan pembelajaran dari evaluasi sebelumnya harus dimasukkan pada tahap alasan dan tujuan, ketika masalah yang akan ditangani sedang dieksplorasi. Evaluator sering mengambil peran aktif dalam pengembangan tujuan yang terdefinisi dengan baik yang menetapkan dengan tepat perubahan apa yang ingin dicapai oleh intervensi dan bagaimana perubahan ini akan diukur.

Pada tahap penilaian, ketika opsi untuk mengatasi masalah sedang diperiksa secara rinci, bukti evaluasi sebelumnya akan sangat berharga untuk menilai kelayakan dan biaya opsi ini. Pemikiran evaluasi awal dan uji coba dapat menjadi sangat penting dalam menguji ide-ide kebijakan (menjelajahi pertanyaan-pertanyaan seperti: apakah ini akan berhasil? mengapa? bagaimana? untuk siapa?).

Bukti evaluasi berguna ketika merancang intervensi baru atau meninjau kebijakan yang ada. Seberapa berguna hal ini terbukti bergantung pada seberapa disesuaikan desain evaluasi dengan kebutuhan pembuat keputusan. Iterasi adalah hal biasa, dan pembelajaran awal dari pemantauan dan evaluasi dapat menghasilkan perubahan cepat pada desain dan tujuan kebijakan.



Desain evaluasi 'Agile' menjadi lebih populer, dengan putaran umpan balik yang cepat untuk mempengaruhi desain intervensi dan pengiriman.

1.3. Waktu Evaluasi

Evaluasi sering dapat dianggap sebagai sesuatu yang terjadi setelah intervensi dilaksanakan. Namun, evaluasi harus menginformasikan pemikiran selama siklus ROAMEF sebelum, selama dan setelah implementasi - dan memiliki kegunaan maksimum jika dipikirkan dengan cara ini.

Sebelum intervensi sepenuhnya terbentuk, evaluasi harus digunakan untuk membantu membentuk desainnya dan bagaimana implementasinya. Menggunakan bukti evaluasi yang ada, bekerja melalui Teori Perubahan (Bab 2), percontohan dan pengujian awal ide kebijakan dapat mengeksplorasi:

- bagaimana intervensi diharapkan bekerja dan bukti apa yang mendukung pemikiran ini;
- mengapa intervensi mungkin tidak berhasil dan bukti apa yang mendukung hal ini;
- di mana letak risiko dan ketidakpastian;
- bagaimana kebijakan bekerja pada skala yang lebih kecil dan dalam lingkungan yang terkendali;
- bukti dasar apa yang harus digunakan untuk mengukur perubahan di masa depan.

Evaluasi Selama Program Berjalan

Evaluasi yang dilaksanakan pada saat program berjalan merupakan kesempatan terbesar bagi evaluasi untuk mempengaruhi keputusan dan membantu memastikan kebijakan dapat mewujudkan manfaat yang dimaksudkan. Selama implementasi, evaluasi biasanya akan melihat perolehan bukti tentang kemandirian desain kebijakan, implementasinya, dan hasil yang muncul. Itu dapat memeriksa pertanyaan-pertanyaan seperti:

- apakah kebijakan disampaikan sebagaimana dimaksud?
- apakah kebijakan berjalan sebagaimana dimaksud?
- apakah ini bekerja dengan cara yang sama di area yang berbeda dan untuk kelompok sasaran yang berbeda?
- apa indikasi awal dari kemungkinan ukuran efek?
- bagaimana cara kerjanya dalam praktik?

- apakah ada konsekuensi yang tidak diinginkan?
- dapatkah desain atau pengiriman ditingkatkan ('penyesuaian dalam penerbangan')?

Setelah Program Selesai

Suatu kebijakan diimplementasikan adalah ketika seluruh kebijakan dapat diperiksa dengan melihat pernyataan yang lebih konklusif tentang desain, implementasi, dan hasil, menjawab pertanyaan seperti sebagai:

- apakah kebijakan itu berhasil?
- apa ukuran efek dan biayanya (dan apakah itu berbeda untuk kelompok yang berbeda)?
- kontribusi apa yang dibuat kebijakan terhadap hasil? • bagaimana hal ini berhubungan dengan apa yang diprediksi dalam penilaian?
- apakah ada dampak yang tidak diinginkan atau negatif?
- apakah ini mewakili nilai-untuk-uang?
- apa yang telah kita pelajari tentang apa yang berhasil di ruang ini? Apa yang bisa dipindahtangankan? pelajaran?

Evaluasi perlu direncanakan sejak dini untuk memastikan bahwa:

- evaluasi memberikan temuan yang berguna bagi mereka yang merancang dan mengimplementasikan kebijakan;
- evaluasi memiliki kualitas yang sesuai untuk tujuan penggunaannya;
- data yang benar dikumpulkan dengan cara yang paling hemat biaya dan seefisien mungkin.

Hubungan antara desain intervensi dan kegunaan evaluasi sangat penting. Perubahan kecil dalam desain intervensi dapat membuat perbedaan antara evaluasi yang bermanfaat dan berkualitas tinggi dengan yang tidak dapat menjawab pertanyaan kunci yang sedang dipertimbangkan (apakah berhasil? sejauh mana? untuk siapa? dan mengapa?). Perubahan desain ini bisa besar – misalnya, dengan mengalokasikan intervensi secara acak untuk membentuk kelompok perlakuan dan kontrol / pembanding seperti yang akan terjadi dengan Uji Coba Kontrol Acak – atau kecil – misalnya memastikan bahwa data administratif tingkat individu tersedia untuk tujuan analisis, atau dengan membuat kuesioner wajib yang harus diisi oleh semua penerima produk atau layanan atau dengan memilih urutan optimal di mana sebuah program diluncurkan.

Sangat penting untuk merencanakan lebih awal data dan

bukti apa yang harus dikumpulkan sebelum implementasi dan selama masa intervensi. Jika kegiatan ini dibiarkan sampai setelah masa intervensi, mungkin membatasi kemampuan untuk melakukan evaluasi yang tepat.

Misalnya, mungkin perlu untuk menemukan kelompok pembanding yang sesuai atau mengumpulkan data dasar sebelum implementasi sehingga kontrafaktual dapat diperkirakan. Hal ini juga terjadi bahwa penerima mungkin lebih setuju untuk diwawancarai saat program ditayangkan dan badan pengiriman mungkin hanya ada selama program berlangsung.

1.4. Jenis Evaluasi

1.4.1. Evaluasi Proses

“Apa yang bisa dipelajari dari bagaimana intervensi itu disampaikan?”

Evaluasi proses cenderung memeriksa kegiatan yang terlibat dalam implementasi intervensi dan jalur di mana kebijakan itu disampaikan. Ini mungkin sangat bervariasi sesuai dengan sifat intervensi dan akan menjadi spesifik kebijakan, yang mencakup pertanyaan-pertanyaan seperti:

- apa yang berhasil dengan baik dan kurang baik, dan mengapa?
- apa yang bisa diperbaiki?
- bagaimana konteks memengaruhi penyampaian?

Evaluasi proses biasanya menggunakan berbagai metode, baik kuantitatif maupun kualitatif yang mencakup isu-isu subjektif (seperti persepsi tentang seberapa baik kebijakan telah beroperasi) dan isu-isu objektif (rincian faktual tentang bagaimana intervensi telah beroperasi, biasanya menggunakan data administratif, jika tersedia).

1.4.2. Evaluasi Dampak

“Apa perbedaan yang dibuat oleh intervensi?”

Evaluasi dampak berfokus pada perubahan yang disebabkan oleh intervensi; pencapaian terukur yang merupakan diri mereka sendiri, atau berkontribusi pada, tujuan intervensi.

Pertanyaan evaluasi dampak yang umum meliputi:

- Hasil terukur apa, baik yang disengaja maupun tidak, yang terjadi?
- Berapa banyak dari hasil ini yang dapat dikaitkan dengan

intervensi?

- Apakah kelompok yang berbeda telah terkena dampak dengan cara yang berbeda, bagaimana dan mengapa?
- Bagaimana konteks mempengaruhi hasil?
- Dapatkah intervensi direproduksi?

1.4.3. Evaluasi *Value-for-money*

“Apakah intervensi ini merupakan penggunaan sumber daya yang baik?”

Sementara dampak menunjukkan dan mengukur hasil, itu tidak dapat menilai sendiri apakah hasil tersebut dapat dibenarkan. Evaluasi nilai-untuk-uang mempertimbangkan isu-isu seperti itu, termasuk apakah manfaat kebijakan lebih besar daripada biayanya, dan apakah intervensi tetap merupakan penggunaan sumber daya yang paling efektif.

Evaluasi dasar nilai-untuk-uang akan membandingkan biaya dan manfaat yang dicapai melalui program dengan harapan awal yang digariskan dalam penilaian, biasanya kasus bisnis atau penilaian dampak. Evaluasi ekonomi yang lebih berkembang juga akan membandingkan manfaat dan biaya dari cara lain untuk mencapai tujuan strategis yang sama.

1.5. Kriteria Disain Evaluasi

Kriteria evaluasi yang baik ditentukan oleh banyak faktor yang berkaitan dengan kebijakan, penggunaan bukti evaluasi yang akan digunakan, dan desain dan pelaksanaan evaluasi itu sendiri. Faktor-faktor ini harus dipertimbangkan sejak awal perencanaan evaluasi dan diperjelas dalam diseminasi temuan evaluasi.

Dalam hal sifat kebijakan dan penggunaan bukti evaluasi, faktor penting yang perlu dipertimbangkan dalam desain evaluasi adalah:

- sifat sistem sebelum intervensi diperkenalkan, yaitu kompleksitas kebijakan yang ada, ukuran populasi, jenis kegagalan pasar, kepentingan pemangku kepentingan, dll;
- skala, kompleksitas, dan tingkat inovasi intervensi;
- bukti yang ada pada kebijakan dan, oleh karena itu, ketidakpastian yang terlibat;
- sifat kebijakan: apa kemungkinan konsekuensinya dan untuk siapa?

- waktu dan sifat keputusan kebijakan utama;
- ruang lingkup untuk mengubah kebijakan saat ini;
- status dan nilai polis.

Evaluasi yang baik akan menjadi evaluasi yang sesuai dengan tujuan: skalanya proporsional dan mencerminkan kebutuhan para pembuat keputusan dan mereka yang meneliti kebijakan dari luar.

Prinsip umum yang dapat memandu keputusan seputar evaluasi untuk mempertahankan kualitas tinggi meliputi:

- **Berguna:** Evaluasi berkualitas tinggi bila dirancang untuk memenuhi kebutuhan banyak pemangku kepentingan yang terlibat. Pemangku kepentingan (Evaluasi) dan menghasilkan keluaran yang berguna dan dapat digunakan pada saat yang tepat. Menyesuaikan evaluasi di sekitar titik keputusan yang diketahui dan area perdebatan kebijakan adalah kunci untuk mencapai tujuan ini. Selain itu, komunikasi yang jelas tentang keterbatasan temuan evaluasi harus selalu diberikan untuk memastikan hasil digunakan secara bertanggung jawab.
- **Kredibel:** Agar bermanfaat, evaluasi harus kredibel. Hal ini sering dicapai dengan memastikan tingkat objektivitas. Hal ini dapat dicapai melalui pekerjaan evaluasi yang dilakukan oleh kelompok evaluator independen, atau melalui tokoh-tokoh independen yang disegani yang mengarahkan dan meninjau ulang baik desain maupun keluaran dari pekerjaan evaluasi. Transparansi sangat penting.
- **Kuat:** Meskipun tidak ada kriteria objektif untuk kualitas, evaluasi harus dirancang dengan baik, dengan pendekatan dan metode evaluasi yang sesuai, dan dilaksanakan dengan baik (misalnya, memastikan strategi pengambilan sampel yang memadai dan ukuran sampel dalam survei untuk mengidentifikasi perubahan yang signifikan secara statistik; mencapai kekuatan yang cukup dalam desain eksperimental; pengambilan sampel kualitatif yang memastikan berbagai suara terdengar; penilaian yang memadai atas ketidakpastian input dan output model; kepatuhan terhadap prinsip-prinsip etika). Pendekatan untuk menetapkan dampak harus melibatkan perbandingan yang ketat baik dalam waktu, antar kelompok atau dengan teori alternatif. Tinjauan sejawat independen dan pengarah independen dapat membantu memastikan kualitas desain dan pelaksanaan evaluasi.

- Proporsional: Proporsionalitas adalah konsep kunci dalam evaluasi. Tidak semua intervensi akan membutuhkan tingkat pengawasan yang sama atau memiliki kebutuhan belajar yang sama. Dalam kasus intervensi berisiko rendah, terbukti dengan baik, dan berprioritas rendah, latihan pemantauan dan evaluasi dengan sentuhan ringan untuk memastikan bahwa intervensi telah disampaikan sebagaimana dimaksud dan mencapai hasil yang diprediksi kemungkinan besar adalah semua yang diperlukan. Di sisi lain, kebijakan berisiko tinggi dan berstatus tinggi yang membuka jalan baru kemungkinan memerlukan evaluasi skala besar.

Kriteria untuk intervensi 'prioritas' yang memerlukan evaluasi substansial biasanya meliputi:

- kebijakan profil tinggi;
- tingkat ketidakpastian/risiko yang tinggi (termasuk kemungkinan konsekuensi negatif);
- biaya tinggi (jika mengevaluasi percontohan, biaya penuh untuk meluncurkan kebijakan harus dipertimbangkan);
- potensi pembelajaran yang tinggi (intervensi prioritas rendah pada kriteria lain dapat memiliki potensi tinggi untuk mengisi kesenjangan bukti yang penting secara strategis).

1.6. Pemangku Kepentingan Evaluasi

Pada intinya, evaluasi harus bermanfaat bagi banyak audiens dan penggunanya yang berbeda. Membangun desain evaluasi di sekitar kebutuhan pengguna akan memastikan mereka terlibat dan bahwa output berguna, dan digunakan oleh, mereka.

Agar evaluasi memenuhi tujuan kembarnya yaitu akuntabilitas dan pembelajaran, ada sejumlah 'pelanggan' berbeda yang kebutuhannya harus dipertimbangkan.

- Mereka yang bertanggung jawab atas intervensi yang sedang dipertimbangkan: ini adalah orang-orang yang paling banyak memperoleh manfaat dari bukti yang dapat mengurangi risiko dan ketidakpastian, dan dari mempelajari apa yang berhasil dan apa yang tidak.
- Mereka yang bertanggung jawab atas kebijakan masa depan: kelompok ini akan membutuhkan bukti tentang apa yang berhasil (dan/atau tidak), mengapa dan

bagaimana, dan tentang pelajaran yang dapat ditransfer.

- Mereka yang bertanggung jawab untuk analisis penilaian: mereka akan memiliki wawasan paling luas tentang bukti dan data apa yang hilang dari penilaian intervensi, dan apa yang akan berguna untuk penilaian kebijakan di masa depan.
- Mereka yang bertanggung jawab untuk meneliti keputusan dan pengeluaran pemerintah: mereka yang meminta pertanggungjawaban pemerintah adalah audiens yang bersemangat untuk mendapatkan bukti seputar kemandirian rancangan dan pelaksanaan intervensi, serta dampak dan biayanya.
- Peserta/ penerima kebijakan: mereka yang terkena dampak kebijakan biasanya juga merupakan peserta kunci dalam evaluasi. Masukan mereka diperlukan, tetapi mereka juga akan memiliki kebutuhan bukti dan perspektif tentang elemen kebijakan apa yang harus difokuskan
- Pihak yang memberikan intervensi: biasanya, meskipun kebijakan sering dirancang di pemerintah pusat, kebijakan tersebut disampaikan oleh pihak lain, dalam banyak kasus melalui rantai pengiriman yang panjang. Evaluasi harus sesuai dengan kebutuhan dan masalah semua pihak dalam rantai pengiriman.
- Publik (sering melalui media): garis utama akuntabilitas adalah kepada publik yang ingin mengetahui bahwa uang pemerintah dibelanjakan dengan bijaksana, dan bahwa kita belajar dari pengalaman masa lalu.
- Akademisi / peneliti lain: pemerintah jarang menjadi satu-satunya pihak yang berkepentingan dalam bidang kebijakan tertentu. Akademisi dan peneliti lain sering menghabiskan waktu untuk meneliti data pemerintah. Penting untuk bekerja dengan mereka untuk memastikan penggunaan terbaik dari bukti penelitian sedang dibuat dan pembelajaran maksimal sedang diekstraksi.

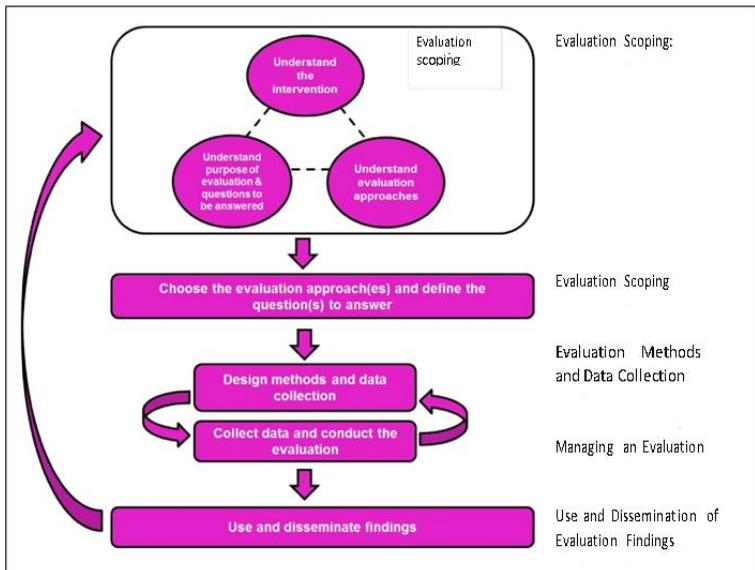
Membangun desain evaluasi di sekitar pengguna layanan, calon pengguna evaluasi dan penggunaan yang mereka inginkan untuk menempatkan bukti evaluasi akan membantu memaksimalkan

kegunaan dan penggunaan temuan evaluasi.

Evaluasi, meskipun konsepnya sederhana – pada dasarnya, belajar dari implementasi – dapat menjadi tugas yang sangat teknis dan kompleks untuk dilakukan dengan baik. Ada komunitas evaluasi global yang besar, dan sejumlah teknik spesialis dan kerangka kerja konseptual yang dapat digunakan. Hasil dari, evaluasi adalah sesuatu yang dirancang, diawasi, dan dikelola dengan baik oleh individu dan tim dengan keahlian khusus.

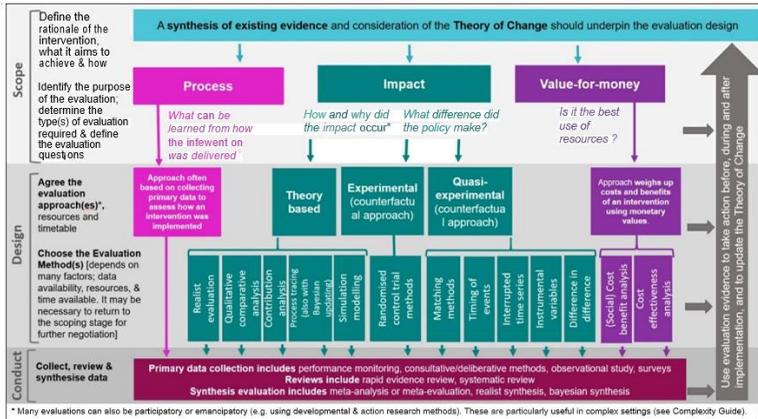
Di pemerintahan, evaluasi biasanya dirancang oleh evaluator spesialis atau analis dari penelitian sosial, ekonomi, statistik dan riset operasional, bekerja sama dengan mereka yang merancang dan mengimplementasikan intervensi. Lebih sering daripada tidak, evaluasi aktual itu sendiri – mengumpulkan data yang ada dan data baru serta menganalisis dan menafsirkan temuan – dikontraskan kepada spesialis independen.

1.7. Tahapan Evaluasi



Gambar 2. Tahapan Utama Evaluasi

BAB 2. RUANG LINGKUP EVALUASI



Gambar 3. Cakupan Evaluasi

Dari Gambar 3 terlihat bahwa cakupan evaluasi meliputi penetapan cakupan evaluasi (*scope*), dilanjutkan dengan penyusunan dan penetapan disain evaluasi (*design*) dan pelaksanaan evaluasi (*conduct*).

2.1. Tahapan Penetapan Cakupan Evaluasi

Tahapan ini terdiri dari 4 (empat) bagian, yakni pertama tahap memahami kembali program yang dilakukan termasuk intervensi yang dilakukan dengan memahami teori perubahan, kedua adalah tahapan penetapan maksud dan tujuan evaluasi, dan ketiga adalah tahap untuk mengidentifikasi jenis evaluasi yang diperlukan, serta tahap keempat adalah memahami pendekatan evaluasi dampak.

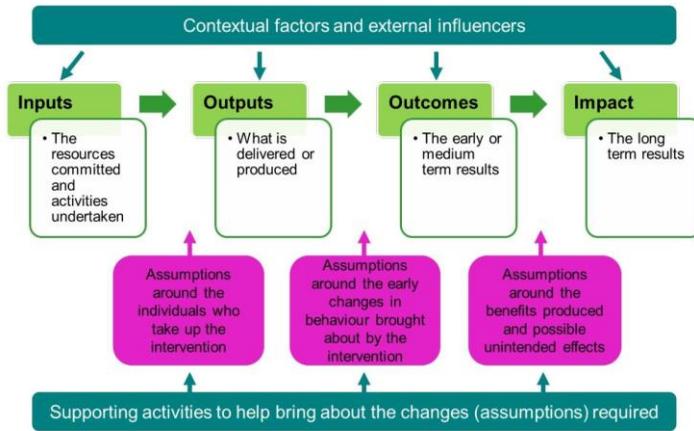
2.2.1. Memahami Program yang Telah Dilakukan

Pembuatan kebijakan yang baik memerlukan pemahaman menyeluruh tentang intervensi dan bagaimana intervensi tersebut diharapkan untuk mencapai hasil yang diharapkan. Evaluasi yang baik juga membutuhkan pemahaman ini. Memeriksa secara menyeluruh intervensi yang diusulkan memastikan:

- pemahaman tentang bagaimana intervensi diharapkan berhasil dalam praktik, misalnya:
- masalah intervensi bertujuan untuk mengatasi; perubahan itu bertujuan untuk membawa; rantai sebab akibat dari peristiwa yang diharapkan membawa perubahan; aktor utama; kelompok yang diperkirakan akan terkena dampak; dan kondisi yang diharapkan yang diperlukan agar intervensi berhasil;
- mengungkap asumsi yang menjadi dasar intervensi dan kekuatan atau kelemahan bukti yang mendukung asumsi ini;
- pemeriksaan konteks yang lebih luas, seperti perubahan kebijakan lain atau perubahan dalam faktor ekonomi, sosial dan lingkungan;
- perancang dan pelaksana intervensi memiliki kesempatan untuk menguji desain intervensi dan memastikan mereka setuju tentang bagaimana intervensi diharapkan bekerja.

Memahami intervensi biasanya dilakukan melalui sintesis bukti yang ada dan menghasilkan Teori Perubahan. Sebuah Teori Perubahan menangkap semua detail yang tercantum di atas, termasuk teori tentang bagaimana intervensi diharapkan bekerja (menetapkan semua langkah yang diharapkan untuk terlibat dalam mencapai hasil yang diinginkan), asumsi yang dibuat, kualitas dan kekuatan intervensi, bukti yang mendukungnya, dan faktor kontekstual yang lebih luas.

Bagian penting dalam menghasilkan Teori Perubahan adalah sintesis dari bukti yang ada. Dengan menyatukan dan menilai kekuatan bukti yang ada tentang intervensi, evaluator dapat mulai melihat di mana bukti lebih lemah. Ini dapat membantu mengidentifikasi pertanyaan kunci yang perlu dijawab oleh evaluasi. Mengembangkan Teori Perubahan biasanya melibatkan pertimbangan masukan yang diusulkan (investasi/peraturan/tindakan apa yang akan dilakukan) dan rantai sebab akibat yang mengarah dari masukan ini hingga keluaran dan hasil yang diharapkan (lihat gambar 2.2). Ini mempertimbangkan mekanisme kausal dimana intervensi diharapkan untuk mencapai hasilnya, mendasarkan teori ini pada pengumpulan dan sintesis bukti.



Gambar 4. Contoh Teori Perubahan Linier

Ada banyak alat pemetaan yang dapat digunakan untuk membantu mengeksplorasi intervensi dan bagaimana intervensi tersebut diharapkan bekerja, sering digambarkan sebagai 'teori program'. Ini termasuk pemetaan Teori Perubahan, pemetaan logika, bingkai log, pemetaan manfaat, dan pemetaan sistem¹⁶. Alat yang paling tepat untuk digunakan akan tergantung pada karakteristik intervensi, kompleksitas sistem yang diterapkan, dan jenis evaluasi yang sedang direncanakan.

Mengembangkan Teori Perubahan biasanya akan melibatkan pemangku kepentingan yang terlibat dalam merancang dan melaksanakan intervensi. Bisa melalui workshop atau konsultasi. Selain itu, metode penelitian, termasuk sintesis bukti, kelompok fokus, dan panel ahli, dapat digunakan untuk mengumpulkan dan mensintesis bukti untuk digunakan dalam pengembangannya.

Teori Perubahan dapat berkisar dari deskripsi sederhana hingga analisis yang lebih kompleks. Latihan yang lebih canggih menghasilkan penilaian yang lebih rinci dan ketat dari intervensi dan asumsi yang mendasarinya. Mereka merinci: mekanisme kausal yang tepat yang memimpin dari satu langkah ke langkah berikutnya; mekanisme alternatif untuk hasil yang sama; asumsi di balik masing-masing langkah kausal; bukti yang mendukung asumsi ini; dan bagaimana faktor kontekstual, perilaku, dan organisasi yang berbeda dapat memengaruhi bagaimana, atau jika, hasil itu terjadi. Hal ini juga berguna untuk mengeksplorasi teori program negatif: semua alasan

mengapa langkah-langkah kausal mungkin tidak terjadi dalam praktik dan mengapa hasil yang diinginkan kemungkinan tidak terjadi. Ini dapat membantu mengidentifikasi risiko dan masalah untuk dijelajahi sebagai bagian dari evaluasi. Dalam intervensi yang lebih kompleks, Teori Perubahan dapat menangkap interaksi, loop umpan balik dan analisis batas sistem.

Cara terbaik untuk mengembangkan Teori Perubahan adalah melalui kolaborasi ekstensif dengan berbagai pemangku kepentingan, termasuk perancang, pelaksana, penerima manfaat, dan/atau kelompok kepentingan untuk memahami bagaimana kebijakan tersebut mungkin berjalan dari berbagai perspektif. Ketika ini tidak mungkin, seminimal mungkin harus diuji stres dengan pemangku kepentingan utama untuk menguji apakah itu mencerminkan pandangan mereka tentang bagaimana kemungkinan intervensi akan berhasil. Jika intervensi besar dengan banyak elemen, serangkaian model dapat dikembangkan dengan bermanfaat, dengan fokus pada berbagai aspek intervensi.

Yang terpenting, Teori Perubahan harus terus dikembangkan selama masa evaluasi karena bukti baru dikembangkan.

2.2.2. Menetapkan Tujuan/ Manfaat Evaluasi

Evaluasi dapat dirancang untuk menjawab berbagai pertanyaan potensial. Penting untuk memperjelas sejak awal apa pertanyaan-pertanyaan ini dan bagaimana temuan dari pertanyaan-pertanyaan itu diharapkan digunakan, oleh siapa dan kapan. Ini akan menginformasikan pendekatan evaluasi yang akan digunakan, membantu memfokuskan evaluasi, dan memastikan temuan memiliki peluang terkuat untuk berdampak pada pengambilan keputusan. Pertanyaan yang akan dijawab oleh evaluasi akan diinformasikan oleh:

- tujuan evaluasi yang dinyatakan;
- pertanyaan yang diidentifikasi oleh Teori Perubahan, misalnya:
 - ✓ Apa saja bidang ketidakpastian?
 - ✓ Di mana kelemahan penting dalam basis bukti?
 - ✓ Apa hasil antara yang dapat mengukur kemajuan menuju hasil akhir? hasil?

- pertanyaan yang ingin dijawab oleh pemangku kepentingan (misalnya pendanaan, perancangan, pelaksanaan atau dampak intervensi.);
- poin keputusan awal untuk intervensi (misalnya apakah akan melanjutkan peluncurannya, apakah akan memodifikasi desain atau implementasinya, dan poin tinjauan yang diharapkan atau direncanakan) dan bukti apa yang diperlukan untuk menginformasikan keputusan tersebut;
- bagaimana temuan diharapkan digunakan, dengan mempertimbangkan kebutuhan jangka pendek (misalnya manfaat realisasi) dan kebutuhan jangka panjang (misalnya menjawab pertanyaan 'apa yang berhasil?' dan 'mengapa?' untuk menginformasikan kebijakan serupa).

Evaluasi dibingkai oleh daftar pertanyaan evaluasi yang harus dijawab. Dalam kombinasi dengan teori program, mereka mendefinisikan ruang lingkup evaluasi; sangat penting untuk keberhasilan evaluasi untuk mendapatkan mereka benar.

Mencakup daftar pertanyaan evaluasi potensial kemungkinan akan menghasilkan daftar yang panjang. Beberapa evaluasi akan dapat menjawab setiap pertanyaan yang diajukan, biasanya karena waktu dan sumber daya yang diperlukan untuk menjawab semuanya, serta keterbatasan metodologis.

2.2.3. Memahami Jenis Evaluasi Yang Diperlukan

Pertanyaan yang dihasilkan dalam latihan pelingkupan kemungkinan mencakup ketiga jenis evaluasi: proses, dampak, dan nilai untuk uang. Jenis pertanyaan yang berhubungan dengan masing-masing disajikan pada Tabel 1.2. Tabel pertanyaan evaluasi. Untuk pemahaman penuh tentang intervensi, aspek dari ketiga jenis evaluasi mungkin diperlukan.

Selalu ada kelemahan dalam setiap evaluasi ini: misalnya, mengetahui sejauh mana dampak suatu intervensi jarang akan menjelaskan mengapa dampak ini terjadi. Ini sangat penting ketika hasil yang diharapkan tidak tercapai. Dalam hal ini, evaluasi proses dapat memberikan bukti penting untuk memahami apakah masalah tersebut merupakan hasil dari rancangan intervensi atau penyampaian intervensi dan apakah masalah yang teridentifikasi ini dapat diatasi.

Tabel 1. Pertanyaan Evaluasi

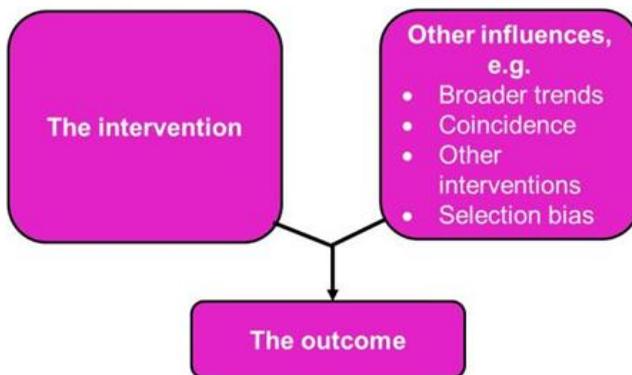
<p>Pertanyaan evaluasi proses: Apa yang dapat dipelajari dari bagaimana intervensi disampaikan?</p>	<p>Pertanyaan evaluasi dampak: Apa perbedaan intervensi?</p>	<p>Pertanyaan evaluasi nilai-untuk-uang: Apakah ini penggunaan sumber daya yang baik?</p>
<p>Apakah intervensi disampaikan sebagaimana dimaksud?</p> <ul style="list-style-type: none"> • Apakah ada cukup sumber daya? • <p>Apakah ada masalah yang tidak diharapkan atau tidak diinginkan dalam penyampaian intervensi?</p> <ul style="list-style-type: none"> • Se jauh mana intervensi telah menjangkau semua orang yang dimaksudkan? <p>Apa yang berhasil dengan baik, atau kurang baik, untuk siapa dan mengapa?</p> <p>Apa yang bisa diperbaiki?</p> <p>Apa yang bisa dipelajari dari metode penyampaian yang digunakan?</p> <ul style="list-style-type: none"> • Mungkinkah intervensi telah dilakukan dan dikirim dengan biaya lebih murah? <p>Bagaimana konteks memengaruhi penyampaian?</p> <ul style="list-style-type: none"> • Bagaimana faktor eksternal mempengaruhi pelaksanaan dan fungsi intervensi? • Bagaimana faktor eksternal mempengaruhi sikap dan perilaku kelompok sasaran? 	<p>Apakah intervensi mencapai hasil yang diharapkan?</p> <ul style="list-style-type: none"> • Se jauh mana? <p>Apakah intervensi menyebabkan perbedaan?</p> <ul style="list-style-type: none"> • Se jauh mana hasil dapat dikaitkan dengan intervensi? Seberapa yakin kita bahwa intervensi menyebabkan perubahan yang diamati? • Faktor penyebab apa yang menyebabkan dampak yang diamati? • Berapa banyak yang dapat dikaitkan dengan faktor eksternal? • Apa yang akan terjadi? <p>Bagaimana konteks memengaruhi hasil?</p> <ul style="list-style-type: none"> • Apakah intervensi menghasilkan hasil yang tidak diinginkan? • Apakah hasil dipengaruhi oleh faktor eksternal lainnya? <p>Se jauh mana kelompok yang berbeda telah terpengaruh dengan cara yang berbeda, bagaimana dan mengapa?</p> <p>Bisakah intervensi direproduksi?</p> <p>Pelajaran umum apa yang telah kita pelajari tentang dampak?</p>	<p>Seberapa efektif biaya intervensi?</p> <ul style="list-style-type: none"> • Biaya per unit (hasil, peserta, dll.) • Berapa biaya untuk memberikan intervensi? <ul style="list-style-type: none"> • Apakah intervensi tersebut efektif dari segi biaya? (dibandingkan dengan alternatif dan dibandingkan dengan tidak melakukan apa-apa)? • Apa pilihan yang paling hemat biaya? <p>Apa nilai-untuk-uang dari intervensi?</p> <ul style="list-style-type: none"> • Apa manfaatnya? • Berapa biayanya? • Apakah manfaatnya lebih besar daripada biayanya? • Berapa rasio biaya terhadap manfaat? <p>Apakah intervensi merupakan penggunaan sumber daya yang terbaik?</p> <ul style="list-style-type: none"> • Bagaimana rasio biaya terhadap manfaat dibandingkan dengan intervensialternatif?
<p>Pembelajaran Masa Depan. Berbagai jenis evaluasi bersama-sama dapat membantu menjawab pertanyaan yang digunakan untuk pembelajaran di masa mendatang:</p> <ul style="list-style-type: none"> • Apakah tujuan intervensi relevan, dalam konteks yang berbeda? • Dapatkah kebijakan tersebut diharapkan berhasil dalam konteks lain? • Apakah intervensi berkelanjutan dari perspektif keuangan, ekonomi, sosial dan lingkungan? • Apa yang telah dipelajari tentang bagaimana mengintervensi ruang intervensi ini yang dapat dialihkan ke inisiatif lain dan penilaian di masa depan? 		

2.2.4. Memahami Pendekatan Evaluasi Dampak

Dalam memilih pendekatan evaluasi yang paling cocok, seorang evaluator harus mempertimbangkan jenis evaluasi dan pertanyaan yang harus dijawab, di samping pemahaman tentang intervensi itu sendiri, konteks di mana intervensi itu dilaksanakan dan informasi/ data yang tersedia.

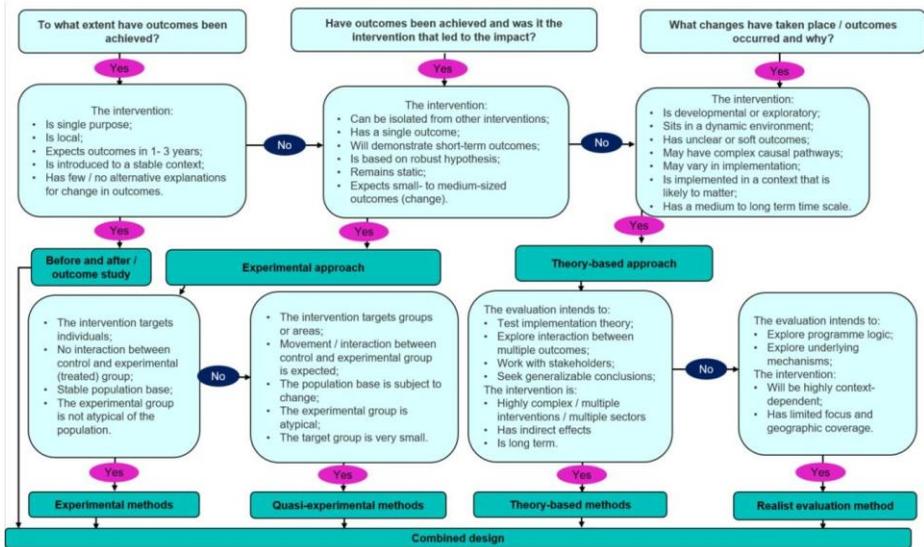
Evaluasi dampak bisa sangat menantang untuk dirancang dan diimplementasikan. Evaluasi dampak bertujuan untuk menilai perubahan apa yang telah terjadi dan skala perubahan tersebut. Mereka juga menilai sejauh mana perubahan dapat dikaitkan dengan intervensi, di atas dan di atas apa yang akan terjadi

jika intervensi tidak terjadi. Ini rumit karena akan ada pengaruh lain yang harus dipahami untuk mengklaim bahwa intervensi telah berpengaruh (Gambar 5).



Gambar 5. Faktor yang mempengaruhi hasil

Tidak ada satu pendekatan evaluasi yang dapat secara tepat mengevaluasi semua jenis intervensi – setiap desain memiliki kelebihan dan kekurangan – dan seringkali pendekatan mungkin perlu digabungkan. Memilih pendekatan evaluasi dampak adalah keputusan awal yang akan mempengaruhi semua langkah selanjutnya.



Gambar 2. Memilih pendekatan untuk evaluasi dampak, berdasarkan pertanyaan yang harus dijawab.

(Direproduksi dari: Hills, D. dan Junge, K. (2010). Panduan untuk evaluasi dampak transportasi: Memilih pendekatan evaluasi untuk mencapai atribusi yang lebih baik. [pdf]. Institut Tavistock berkonsultasi dengan AECOM. Tersedia di: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/525806/transport-impact-evaluations.pdf

Tabel 2. Memilih pendekatan eksperimental atau kuasi-eksperimental untuk evaluasi dampak

	MORE FEASIBLE IF:	LESS FEASIBLE IF:
Nature of intervention	<ul style="list-style-type: none"> The intervention is discrete (can be disentangled from other programme interventions) and is stable The system the intervention is being applied to is relatively stable and unchanging 	<ul style="list-style-type: none"> The intervention is difficult to define or disentangle from other programme interventions or local context The intervention adapts over time. The intervention is being applied to a complex and emergent system (see Supplementary Guidance on evaluating Complexity for more detail)
Nature of impact	<ul style="list-style-type: none"> There is a direct, linear relationship between the expected outcome and the intervention A large effect relative to other changes is expected The effect is realised within a short time period (and does not vanish immediately thereafter) 	<ul style="list-style-type: none"> There is a complex¹ or distant relationship between the expected outcome and the intervention, with many potential confounding factors A small effect is expected The effect builds up gradually over an extended time period The exact nature of the impact is unknown
Data availability: what was done where, when, to whom	<ul style="list-style-type: none"> The intervention involves a distinct change in practice with respect to identifiable participants (individuals, groups, institutions or areas) Data is available on the participants of the intervention Data is available on precise time periods Data to support the evaluation collected before and during the intervention Data can be collected from samples of sufficient size 	<ul style="list-style-type: none"> The intervention involves a consolidation of existing good practice or is poorly differentiated between participants Data is only available as coarsely aggregated totals There is uncertainty over timing of implementation (requires aggregation over time) Data to support evaluation is not sought until the policy is already established, or is unavailable for non-participants Sample sizes are too small
Potential comparison on groups	<ul style="list-style-type: none"> Intervention is built into policy design, and so comparison groups are allocated and data collected from both There is a phased start Random allocation is possible Other objective allocation is possible, for example, using a cut-off score 	<ul style="list-style-type: none"> Intervention is not built into policy design, or data is available only for the pilot areas themselves There is a simultaneous launch nationwide There is subjective allocation, so the comparison and target group are different from the outset

Pendekatan eksperimental dan kuasi-eksperimental cenderung paling cocok jika:

- ada fokus pada akuntabilitas. Pendekatan ini biasanya menghasilkan bukti kuantitatif yang paling dapat dipertahankan tentang dampak;
- hasil yang diharapkan diketahui dan dapat diukur;
- intervensi mempengaruhi sejumlah besar orang/kelompok, tetapi tidak secara keseluruhan populasi;
- intervensi tidak melibatkan sejumlah kegiatan yang berbeda atau implementasi yang bervariasi;
- intervensi diharapkan bekerja dengan cara yang sama untuk kelompok yang berbeda.

Biasanya, pendekatan eksperimental tidak kuat dalam mengidentifikasi dampak yang berbeda pada sub-kelompok tertentu. Pendekatan eksperimental dan kuasi-eksperimental dipilih dengan tujuan utama menilai dampak bersih dari suatu intervensi. Tapi ini sendiri tidak akan menghasilkan wawasan tentang bagaimana perubahan terukur terjadi, atau apakah hasil yang sama akan terjadi jika intervensi dicoba dalam konteks lain atau pada konteks yang berbeda. Menggabungkan pendekatan eksperimental / kuasi-eksperimental dengan pendekatan berbasis teori atau melengkapi dengan bukti evaluasi proses dapat memberikan wawasan yang sering penting ini.

Pendekatan berbasis teori untuk evaluasi dampak

Evaluasi dampak berbasis teori menarik kesimpulan tentang dampak intervensi melalui pengujian ketat apakah rantai sebab akibat yang dianggap membawa perubahan didukung oleh bukti yang cukup kuat dan bahwa penjelasan alternatif dapat dikesampingkan. Evaluasi berbasis teori secara eksplisit berkaitan dengan tingkat perubahan dan mengapa perubahan terjadi; ia mencoba masuk ke dalam kotak hitam tentang apa yang terjadi antara masukan dan hasil, dan bagaimana hal itu dipengaruhi oleh konteks yang lebih luas.

Evaluasi berbasis teori dipusatkan pada Teori Perubahan yang terdefinisi dengan baik, yang mencakup teori tentang penjelasan alternatif untuk hasil. Setelah teori ditetapkan, teori tersebut diuji melalui berbagai sumber bukti.

Saat melakukan pendekatan berbasis teori, penting untuk tetap berpikiran terbuka tentang bukti apa yang dapat digunakan. Literatur eksternal, pendapat ahli, pernyataan publik, penelitian metode campuran dengan berbagai pemangku kepentingan dan pemodelan pertukaran biaya/ manfaat pribadi semuanya dapat digunakan. Ada efisiensi dalam pengumpulan data, dengan memfokuskan upaya pada pengumpulan bukti yang berorientasi pada dampak.

Kekakuan dalam metode berbasis teori berasal dari:

- koherensi teori;
- bukti yang cukup spesifik untuk menguji teori;
- triangulasi berbagai sumber;
- mengesampingkan penyebab alternatif untuk mengklaim dampak; refleksi kritis dan keterbukaan terhadap peer review dan pengawasan eksternal.

Metode evaluasi dampak berbasis teori sangat cocok jika satu atau lebih karakteristik di bawah ini benar:

- ada lanskap kebijakan yang rumit dengan kombinasi intervensi;
- intervensi dirancang untuk membuat perubahan dalam sistem yang kompleks atau di mana ada manajemen adaptif/perubahan intervensi
- hasil yang muncul dan tidak dapat diprediksi sejak awal;
- tidak ada kemampuan untuk mengembangkan kontrafaktual yang sesuai;
- ada persyaratan untuk memahami jika hasil yang sama akan dicapai di tempat atau konteks yang berbeda.

Biasanya, pendekatan-pendekatan ini tidak menghasilkan satu perkiraan numerik dampak, meskipun kuantifikasi dampak mungkin dilakukan, dengan metode yang dipilih yang sesuai.

Pendekatan yang muncul

Pendekatan evaluasi yang lebih baru sedang dikembangkan untuk menyediakan lebih banyak alat untuk digunakan dalam evaluasi kompleksitas, termasuk pendekatan partisipatif. Semakin pendekatan umum adalah evaluasi perkembangan (Patton, 20064), yang didasarkan pada penanganan hasil yang tidak pasti atau muncul dan dimaksudkan untuk menjadi bagian dari lingkaran berkelanjutan dari pengembangan dan iterasi program.

Pendekatan evaluasi perkembangan sering diterapkan pada konteks inovasi, di mana hal itu dapat membantu untuk membingkai pemikiran, memunculkan masalah, dan memberikan umpan balik yang cepat tentang implementasi intervensi.

Ada sejumlah pendekatan alternatif lain yang muncul untuk menangani kompleksitas. Lihat Panduan Tambahan Buku Magenta tentang Mengevaluasi Kompleksitas untuk informasi lebih lanjut.

Pengamatan langsung terhadap dampak

Dalam beberapa kasus yang sangat sederhana, cara intervensi menciptakan perubahan mungkin cukup transparan sehingga dampaknya dapat diamati secara langsung, atau melalui evaluasi proses, tanpa perlu memperhitungkan pengaruh lain. Misalnya, dengan proyek untuk memasok air ke desa di negara berkembang, setiap penurunan yang teramati dalam waktu rata-rata yang dihabiskan anggota rumah tangga untuk mengumpulkan air dapat dikaitkan dengan proyek tersebut tanpa memerlukan kelompok pembanding. Pendekatan ini hanya boleh digunakan jika ada keyakinan bahwa jika intervensi tidak terjadi, tidak ada hal lain yang akan berubah.

2.2. Merancang Evaluasi

Merancang evaluasi melibatkan pengambilan kesimpulan dari tahap pelingkupan dan menyetujui desain evaluasi dan bagaimana evaluasi akan dilakukan. Ini melibatkan pengambilan keputusan pragmatis tentang prioritas bukti dalam kerangka waktu, sumber daya dan kelayakan metode.

Sebuah desain evaluasi harus dipandu oleh prinsip-prinsip yang diuraikan sebelumnya, yaitu kegunaan, kredibilitas, ketahanan dan proporsionalitas.

Merancang evaluasi adalah proses berulang; keputusan tentang metode akan berdampak pada keputusan sebelumnya di mana pertanyaan dapat dan tidak dapat dijawab. Dengan setiap iterasi, desain menjadi lebih detail. Desain evaluasi awal harus memberikan detail yang cukup untuk membuat keputusan tentang pertanyaan evaluasi mana yang layak untuk dijawab dan metode mana yang diperlukan untuk melakukannya. Penilaian kelayakan evaluasi formal adalah tahap desain sementara yang umum, kadang-kadang disebut 'penilaian evaluasi'.

Desain evaluasi akhir spesifik tentang bagaimana metode evaluasi, penelitian dan analisis akan menjawab pertanyaan evaluasi, menguji logika intervensi dan melaporkan garis waktu untuk mendukung titik keputusan yang diidentifikasi.

Rincian ini biasanya ditetapkan dalam 'rencana' atau 'strategi' evaluasi, yang mencakup informasi seperti:

- tujuan intervensi;
- Teori Perubahan;
- rencana untuk tinjauan / sintesis literatur apa pun yang diperlukan; pertanyaan evaluasi dan kapan dan dari siapa buktinya diperlukan;
- bagaimana hasil kepentingan akan diidentifikasi dan diukur;
- bagaimana hasil yang tidak diinginkan akan terdeteksi;
- bagaimana faktor kontekstual yang lebih luas akan diperhitungkan; detail tentang metode yang akan digunakan / bagaimana mereka menjawab evaluasi pertanyaan;
- detail tentang data yang dibutuhkan, bagaimana data itu akan dikumpulkan dan oleh siapa;
- persyaratan sumber daya (pendanaan, staf, keterampilan);
- gambaran umum tentang bagaimana dan kapan bukti evaluasi akan disebarluaskan dan digunakan untuk menginformasikan pembuatan kebijakan.

Saat rencana evaluasi berkembang, keputusan perlu dibuat tentang metode dan pengumpulan data yang diperlukan untuk menjawab pertanyaan evaluasi.

BAB 3.

METODE EVALUASI

Bagian ini mencakup pertimbangan yang diperlukan untuk menyelesaikan desain evaluasi; yaitu, pemilihan metode evaluasi analitik yang paling menjawab pertanyaan evaluasi (berbeda dengan metode evaluasi pengumpulan data yang dibahas selanjutnya). Di samping pertanyaan evaluasi, tiga pertimbangan lebih lanjut penting: proporsionalitas upaya dan sumber daya yang diperlukan untuk menjawabnya; ketersediaan data; dan waktu poin keputusan.

Sebagaimana dinyatakan dalam bab sebelumnya, iterasi akan diperlukan untuk menyempurnakan apa yang mungkin dan praktis. Ini mungkin memerlukan perubahan pada ruang lingkup, waktu atau sumber daya, atau bahkan modifikasi pertanyaan evaluasi. Metode yang berbeda akan memiliki implikasi sumber daya yang berbeda. Setiap pertanyaan evaluasi yang tidak diprioritaskan harus dikomunikasikan dengan jelas dan solusinya disepakati dengan para pemangku kepentingan.

Beberapa metode akan dapat menjawab lebih banyak pertanyaan daripada yang lain dan beberapa akan lebih hemat biaya dan lebih mudah diterapkan; semua memiliki pro dan kontra.

3.1. Memilih Metode yang Tepat

Pemilihan dan penerapan metode analisis untuk evaluasi harus diinformasikan oleh Teori Perubahan dan ketidakpastian serta asumsi yang diidentifikasi. Penting untuk dicatat bahwa berbagai pemangku kepentingan mungkin memiliki

kebutuhan yang berbeda dan memprioritaskan hasil yang berbeda. Hasil ini harus tercermin dalam Teori Perubahan dan pemangku kepentingan harus dilibatkan dalam pemilihan metode evaluasi untuk menyepakati prioritas, mengelola harapan, dan menerima implikasinya.

Harus ada keselarasan antara pertanyaan evaluasi, pendekatan evaluasi dan metode yang digunakan. Dalam praktiknya, sebagian besar desain evaluasi melibatkan metode

campuran, menggabungkan metode kualitatif dan kuantitatif untuk menjawab pertanyaan tentang dampak, proses, dan nilai uang.

Tidak ada metode individual yang dapat memberikan jawaban atas semua pertanyaan evaluasi. Jika waktu dan sumber daya terbatas, pertanyaan dan

metode harus diprioritaskan dan pertukaran antar metode akan diperlukan. Iterasi akan diperlukan.

Penting untuk mempertimbangkan kelayakan dan kelayakan metode yang diusulkan. Biaya, waktu, beban responden, etika, kemungkinan tingkat respons, dan efek potensial dari pengumpulan data itu sendiri terhadap intervensi harus dipertimbangkan.

Populasi yang akan dipelajari dan pertanyaan evaluasi yang harus dijawab harus jelas dari tahap pelingkupan. Definisi yang jelas harus dikembangkan di sekitar individu/kelompok, tempat dan periode waktu yang terkait dengan pertanyaan evaluasi untuk memastikan metode yang paling tepat diidentifikasi. Definisi yang jelas juga dapat membantu memperjelas sejauh mana temuan dapat digeneralisasi.

Waktu pengumpulan data perlu dipertimbangkan secara hati-hati sejalan dengan kapan peristiwa mungkin terjadi dan titik keputusan diidentifikasi selama fase pelingkupan.

3.2. Metode yang digunakan dalam proses dan evaluasi dampak

Metode penelitian yang dirangkum dalam bagian ini biasanya digunakan dalam evaluasi proses dan dampak (dan banyak digunakan dalam penelitian non-evaluatif). Ketika digunakan dalam evaluasi dampak, data yang konsisten juga harus dikumpulkan untuk setiap kelompok kontrol/ pembanding.

Metode penelitian yang biasa digunakan untuk evaluasi diuraikan dalam Tabel 3 di bawah ini. Aspek pengumpulan data dari metode ini tercakup dalam bagian berikutnya. Panduan rinci tentang evaluasi proses dan metodenya dapat ditemukan di tempat lain. Panduan tentang kualitas disediakan dalam Panduan Tambahan, 'Kualitas dalam Evaluasi Kualitatif'.

Tabel 3. Panduan Singkat Metode Penelitian yang biasa digunakan untuk evaluasi

Evaluasi Metode	Keterangan	Pro dan kontra
Wawancara dan fokus kelompok	Wawancara memungkinkan eksplorasi mendalam tentang intervensi dengan peserta. Kelompok fokus berguna untuk mendapatkan pandangan dari kelompok orang daripada individu.	<p>Dapat digunakan untuk memperoleh pandangan individu yang terlibat dalam intervensi.</p> <p>Dapat digunakan untuk mengumpulkan wawasan mendalam tentang intervensi dan menjelaskan pola yang muncul dalam bukti lain yang dikumpulkan (seperti data pemantauan kuantitatif).</p> <p>Dapat menjadi sumber daya yang intensif; membutuhkan waktu untuk melakukan dan menganalisis; tidak memberikan perkiraan numerik; mungkin ada risiko bias dalam pandangan yang dikumpulkan.</p>
Studi kasus	<p>Penyelidikan mendalam terhadap seseorang, kelompok atau peristiwa dalam konteks dunia nyata. Subjek sering dipilih secara purposive karena tidak biasa dan mengungkapkan informasi.</p> <p>Sering menggunakan berbagai sumber bukti dan metode pengumpulan data.</p> <p>Bisa deskriptif, eksploratif atau eksplanatori.</p>	<p>Dapat menangkap situasi kehidupan nyata secara mendalam dan detail dan membantu memahami fenomena yang kompleks.</p> <p>Bekerja dengan baik dalam kombinasi dengan atau melengkapi metode lain, seperti survei.</p> <p>Dapat membantu untuk mengkomunikasikan kepada pemangku kepentingan intervensi apa yang telah berhasil untuk organisasi tertentu dalam konteks tertentu.</p> <p>Sulit untuk menggeneralisasi temuan ke konteks, situasi atau fenomena yang berbeda.</p>
Survei	<p>Biasanya digunakan untuk mengumpulkan data dari sejumlah individu, seperti penerima manfaat, atau organisasi besar dengan banyak anggota staf. Mereka bisa menjadi</p> <p>diberikan secara tatap muka, melalui pos, online, melalui telepon atau sebagai selebaran.</p>	<p>Metode yang efektif untuk memperoleh informasi dari sejumlah besar peserta. Menyediakan data yang cocok untuk analisis statistik yang, jika dirancang dan dilakukan dengan benar, dapat digeneralisasikan ke seluruh populasi yang diinginkan.</p> <p>Kurang berguna untuk memberikan wawasan mendalam tentang intervensi. Mungkin ada masalah tingkat respons yang menurunkan kualitas temuannya.</p>
Pemantauan keluaran atau kinerja	<p>Pengukuran berkelanjutan dan tinjauan kinerja dari suatu intervensi. Rencana pemantauan dikembangkan berdasarkan Teori Perubahan untuk memungkinkan pelacakan input, output, dan hasil intervensi.</p> <p>Untuk meminimalkan kesalahan dalam pengumpulan data dan membebani staf disarankan untuk merancang pemantauan bekerja sama dengan mereka yang akan mengumpulkan data.</p>	<p>Dapat memberikan biaya yang relatif rendah dan metode yang cepat untuk mengidentifikasi apakah suatu intervensi sedang disampaikan dan menciptakan keluaran sebagaimana dimaksud. Menghubungkan dengan baik ke manajemen manfaat.</p> <p>Dapat terasa berat bagi peserta maupun staf (biasanya memberikan intervensi) yang mengumpulkannya.</p>
Studi observasional (termasuk etnografi)	<p>Ini melibatkan mengamati dan mencatat perilaku peserta (termasuk staf penerima intervensi). Sering dilengkapi dengan wawancara dari individu dalam kebiasaan mereka</p> <p>lingkungan untuk menentukan dampak intervensi terhadap kehidupan individu sehari-hari</p>	<p>Dapat memungkinkan pemahaman yang lebih dalam tentang pengalaman individu dari suatu intervensi.</p> <p>Pengamatan dapat membantu meningkatkan akurasi data lain dengan mengurangi bias yang timbul dari petaporan diri oleh peserta. Namun, peserta masih dapat bertindak berbeda jika mereka tahu mereka sedang diamati ('efek Hawthorn'), yang dapat mempengaruhi keakuratan data.</p>

3.3. Metode evaluasi dampak berbasis teori

Metode berbasis teori dapat digunakan untuk menyelidiki dampak bersih dengan menjelajahi rantai sebab akibat yang dianggap membawa

perubahan melalui intervensi. Namun, mereka tidak memberikan perkiraan ukuran efek yang tepat. Evaluasi berbasis teori secara eksplisit berkaitan dengan tingkat perubahan dan mengapa perubahan itu terjadi. Selain itu, seringkali mempertimbangkan konteks pada saat intervensi dilaksanakan.

Metode berbasis teori cenderung sangat cocok untuk evaluasi intervensi kompleks atau intervensi sederhana di lingkungan yang kompleks. Dalam situasi ini, di mana menentukan ukuran efek seringkali sulit, metode berbasis teori dapat mengkonfirmasi apakah intervensi memiliki efek ke arah yang diinginkan. Untuk banyak dari metode ini, tujuannya bukan untuk memberikan bukti definitif bahwa keseluruhan dari setiap perubahan terukur dapat dikaitkan dengan intervensi.

Sebaliknya, mereka bertujuan untuk mengeksplorasi apakah intervensi secara definitif berkontribusi pada perubahan yang terukur. Mereka juga dapat menjelaskan mengapa suatu intervensi berhasil, atau tidak, dan menginformasikan terjemahan ke populasi, tempat, atau periode waktu lain.

Semua metode evaluasi dapat dipertimbangkan dan digunakan sebagai bagian dari pendekatan berbasis teori (khususnya metode penelitian yang bertujuan untuk menjawab pertanyaan evaluasi proses), tetapi yang dijelaskan di bawah ini terutama terkait dengan pendekatan ini.

Sebagai contoh, Evaluasi Realis berusaha mengidentifikasi mekanisme, seringkali psikologis, yang mengubah perilaku manusia sebagai akibat dari intervensi, dengan

mempertimbangkan konteks di mana intervensi terjadi. Evaluasi Realis biasanya bertanya: "apa yang berhasil, untuk siapa, dalam hal apa, bagaimana, dan dalam keadaan apa?". Ini mengembangkan dan menguji satu set hipotesis (atau teori) tentang faktor atau proses yang menjelaskan mengapa intervensi memiliki hasil tertentu (disebut mekanisme), dan apa pengaruh konteks intervensi terhadap mekanisme ini. Mekanisme dapat didefinisikan sebagai menangkap "alasan dan pilihan masyarakat ketika dihadapkan pada suatu ukuran kebijakan". Metode ini dijelaskan secara rinci dalam Panduan Tambahan, 'Evaluasi Realis'.

Pemodelan simulasi tidak secara formal dianggap sebagai metode berbasis teori tetapi dimasukkan di sini karena dapat digunakan secara kuantitatif untuk mewakili skenario kompleks dan memodelkan Teori Perubahan, dampak intervensi dan hasil yang tidak teramati. Jenis simulasi untuk sistem dinamis meliputi: simulasi kejadian diskrit, dinamika sistem, dan pemodelan berbasis agen. Kekuatan model tergantung pada kualitas input dan logika simulasi, yang seringkali memiliki kekurangan. Pemodelan simulasi juga dapat digunakan

untuk menghasilkan kontrafaktual virtual.

Metode berbasis teori bersifat spesialis dan biasanya memerlukan saran ahli untuk membantu mengembangkan pemikiran dan memutuskan metode yang paling tepat. Alat yang berguna dan panduan untuk menginformasikan pemikiran dapat terlihat pada Tabel 4 yang memberikan gambaran umum dan pro dan kontra dari metode berbasis teori yang paling umum.

3.4. Metode evaluasi dampak eksperimental dan kuasi-eksperimental

Metode evaluasi eksperimental dan kuasi-eksperimental digunakan untuk mengukur dampak. Metode ini adalah bahwa ada 'kontrafaktual': hasil yang diamati dari kelompok 'kontrol' yang tidak menerima intervensi, yang dapat dibandingkan dengan hasil dari kelompok intervensi. Kelompok intervensi dan kontrol sama-sama identik secara efektif (biasanya, melalui pengacakan dalam desain eksperimental), atau berbeda dalam cara yang diketahui yang dapat diperhitungkan secara analitis (dalam desain kuasi-eksperimental). Kelompok-kelompok tersebut terdiri dari peserta, yang dapat berupa orang perorangan, atau unit lain, seperti sekolah,

Tabel 4. Panduan untuk Metode Evaluasi Dampak Berbasis Teori Umum bisnis, rumah, atau area spasial. Pencampuran antar kelompok harus minimal untuk membatasi bias dari 'kontaminasi'.

Evaluasi metode	Keterangan	Pro dan kontra
Evaluasi realists	Spesifik, hipotesis kausal 'mekanisme' untuk 'hasil' diartikulasikan dalam 'konteks' dan bukti dikumpulkan untuk masing-masing. 'Mekanisme' menjelaskan mengapa peserta dapat memanfaatkan peluang atau tidak tergantung pada 'konteks', dan pemahaman mereka adalah kunci untuk inferensi kausal.	Menyempurnakan teori dapat mengidentifikasi mekanisme kausal. Dapat menginformasikan dampak jika kontraktual tidak memungkinkan. Memakan waktu, sumber daya intensif dan membutuhkan keahlian subjek-materi. Seringkali sulit untuk dikomunikasikan/interpretasikan karena kerumitannya. Tidak sering memberikan ukuran efek kuantitatif.
Analisis kontribusi	Proses langkah-demi-langkah yang digunakan untuk memeriksa apakah suatu intervensi telah berkontribusi pada hasil yang diamati dengan mengeksplorasi berbagai bukti untuk Teori Perubahan. Ini memberikan garis penalaran yang dibuktikan daripada bukti definitif.	Klaim kontribusi tergantung pada kualitas pemikiran tentang masalah atribusi dan Teori Perubahan. Bekerja pada efek rata-rata – tidak digunakan jika ada variabilitas besar dalam implementasi atau hasil.
Pelacakan proses	Metode terstruktur yang memeriksa satu kasus perubahan untuk menguji apakah mekanisme kausal yang dihipotesiskan, seperti yang diusulkan oleh Theory of Change, menjelaskan hasilnya.	Dapat menguji hipotesis kausal post-hoc. Harus digunakan dengan teliti untuk menegakkan kesalahan inferensial; penjelasan alternatif harus dipertimbangkan dengan cermat. Dukungan untuk satu mekanisme kausal mungkin tidak menghalangi yang lain.
Pembaruan Bayesian	Ditambahkan ke metode berbasis teori lainnya untuk menilai lebih ketat apakah bukti mendukung klaim kontribusi. Probabilitas sejumlah kecil klaim kontribusi diestimasi sebelum observasi kemudian diuji.	Membutuhkan fasilitas yang sangat terjangkau.
Pelacakan kontribusi	Metode campuran partisipatif untuk menetapkan validitas klaim kontribusi dengan kriteria eksplisit untuk memandu evaluator dalam pengumpulan data dan pembaruan Bayesian untuk mengukur tingkat kepercayaan dalam klaim. Termasuk kontribusi 'percobaan' dengan semua pemangku kepentingan untuk menetapkan apa yang akan membuktikan/mendak klaim.	Efisien berfokus pada bukti yang dapat meningkatkan kepercayaan dalam klaim. Meminimalkan bias konfirmasi menggunakan 'teman kritis' dalam fase pengujian. Intervensi membutuhkan waktu untuk memiliki efek yang dapat dideteksi. Harus mengeksplorasi penyebab potensial lainnya. Bukan untuk membandingkan intervensi.
Kualitatif Komparatif Analisis	Digunakan untuk membandingkan beberapa kasus dan secara sistematis memahami pola karakteristik yang terkait dengan hasil yang diinginkan atau tidak diinginkan berdasarkan pengetahuan kualitatif. Dapat menjelaskan penyebab kompleks (kombinasi faktor) dan 'keseimbangan' (beberapa penyebab hasil).	Dapat mengidentifikasi kelompok faktor penyebab dalam evaluasi post-hoc. Secara sistematis menganalisis bukti studi kasus. Bekerja paling baik dengan 10-50 kasus. Membutuhkan data yang konsisten tentang bagaimana faktor-faktor tersebut mempengaruhi hasil dan penilaian yang lebih berhasil di seluruh studi kasus.
panen hasil	Mengumpulkan bukti perubahan dan kemudian bekerja kembali untuk menilai kontribusi terhadap perubahan. Mendorong partisipasi pemangku kepentingan secara real-time untuk pemantauan berkelanjutan.	Berguna di mana partisipasi mudah dan pemangku kepentingan berbeda karena membantu memberikan ketelasan kepada semua. Ini intensif sumber daya.
Perubahan paling signifikan	Metode partisipatif untuk evaluasi dampak intervensi kompleks. Melibatkan pengumpulan cerita perubahan signifikan dari lapangan dan pemilihan sistematis yang paling signifikan oleh panel pemangku kepentingan. Intervensi seringkali juga bersifat partisipatif.	Berguna ketika tidak mungkin untuk memprediksi hasil atau ketika prioritas hasil tidak dapat disepakati. Membangun pemahaman di seluruh pemangku kepentingan. Memakan waktu dan sumber daya yang intensif dan membutuhkan fasilitas yang kuat.

Mengumpulkan dan menganalisis data yang sebanding dari intervensi dan kontrol memungkinkan evaluator untuk secara percaya diri mengaitkan

setiap perubahan terukur pada intervensi (tergantung pada asumsi khusus untuk metode ini). Metode evaluasi dampak ini disukai ketika kita perlu mengetahui rata-rata tambahan atau perubahan bersih yang disebabkan oleh suatu intervensi atau seberapa banyak hasil yang diamati dapat dikaitkan dengan intervensi.

Namun, metode yang tepat untuk digunakan tergantung pada: apakah partisipan dalam suatu intervensi dapat diacak; ukuran efek yang diharapkan; ketersediaan data; dan ketersediaan kelompok kontrol potensial. Cara sederhana, namun efektif, untuk membantu merencanakan atau meninjau metode eksperimental adalah PICOT: kerangka kerja, yang mempertimbangkan populasi, intervensi, kelompok kontrol, hasil dan periode waktu.

Randomized Controlled Trials (RCTs) secara prospektif dan acak mengalokasikan peserta untuk intervensi atau kelompok kontrol. RCT dapat dengan kuat mengevaluasi dampak intervensi karena memperhitungkan faktor yang diketahui dan tidak diketahui karena alokasi untuk pengobatan bersifat acak. Perbedaan terukur antara kelompok dapat, oleh karena itu, dianggap sebagai hasil dari intervensi saja. Namun, kebutuhan untuk memastikan implementasi intervensi yang kaku (kesetiaan tinggi dengan protokol) dapat mengurangi validitas eksternalnya.

Ada sejumlah variasi RCT untuk berbagai kebutuhan dan situasi: RCT faktorial secara independen mengacak peserta ke beberapa intervensi; RCT cluster mengacak kelompok peserta daripada individu; dan RCT step-wedge menerapkan intervensi secara berurutan dan acak pada kelompok peserta.

Metode, seperti *Sequential Multiple Assignment Randomized Trial (SMART)* dan *Multiphase Optimization Strategy (MOST)*, telah dikembangkan untuk optimalisasi intervensi dan mungkin sangat tepat untuk mengembangkan intervensi digital yang lebih kuat.

Metode kuasi-eksperimental kontrafaktual, tetapi tidak ada yang dicapai melalui pengacakan. Selain itu, beberapa metode menjelaskan perbedaan yang diketahui antara kelompok intervensi dan kelompok kontrol dalam analisis. Pemilihan metode biasanya tergantung pada: sifat alokasi ke dalam kelompok intervensi dan kontrol (misalnya, dapat dikendalikan oleh evaluator atau hasil dari kejutan eksternal); sifat kelompok kontrol (bisa bersamaan atau historis); format data yang tersedia (seperti diskrit atau tren); dan volume data yang tersedia.

3.5. Metode Evaluasi *Value-for-Money*

Metode evaluasi nilai-untuk-uang membandingkan manfaat dengan biaya intervensi, termasuk aspek yang merugikan dan tidak diinginkan. Dua metode

yang banyak digunakan adalah analisis efektivitas biaya sosial dan analisis biaya-manfaat sosial, yang keduanya memungkinkan untuk membandingkan dua atau lebih pilihan alternatif (intervensi).

Perbedaan utama antara kedua jenis evaluasi ini adalah:

- analisis efektivitas biaya sosial membandingkan biaya cara-cara alternatif menghasilkan output yang sama atau serupa.
- Analisis biaya-manfaat sosial melangkah lebih jauh untuk menilai dampak dari perbedaan intervensi pada kesejahteraan sosial dengan semua biaya dan manfaat yang relevan yang dinilai dalam istilah moneter (jika proporsional dan memungkinkan).

Estimasi terukur dari keluaran atau dampak intervensi alternatif biasanya berasal dari metode eksperimen, eksperimen semu atau sintesis. Tanpa perkiraan terukur dari efek bersih, metode berbasis teori dapat digunakan untuk menilai apakah dampak tersebut kemungkinan besar akan konsisten dengan titik impas.

Biaya intervensi alternatif yang dipertimbangkan harus dieksplorasi secara rinci untuk menunjukkan mana yang akan memberikan pengembalian investasi yang lebih besar. Biaya ini harus mencakup nilai sosial dan biaya sosial, seperti pekerjaan, kesehatan, kesejahteraan, dan produktivitas jika memungkinkan. Analisis distribusi harus digunakan untuk mempertimbangkan biaya dan manfaat pada kelompok populasi yang berbeda.

Evaluasi metode	Keterangan	Pro dan kontra
Analisis efektivitas biaya sosial	Mendukung pemilihan opsi intervensi yang memiliki ukuran hasil yang sama dengan membandingkan biaya dan ukuran efeknya.	Memungkinkan perbandingan di mana manfaat tidak dapat dikonversi ke dalam unit moneter atau biaya untuk melakukannya adalah penghalang. Secara sistematis menyediakan komparabilitas opsi. Kurangnya manfaat dalam metrik moneter membatasi nilainya dibandingkan dengan analisis biaya-manfaat.
Analisis biaya manfaat sosial	Mengukur dan memonetisasi biaya dan manfaat intervensi memungkinkan perbandingan di antara mereka, bahkan ketika mereka memiliki ukuran hasil yang berbeda. Dapat digunakan untuk meramalkan menggunakan perkiraan biaya dan manfaat atau digunakan pasca-intervensi dalam evaluasi menggunakan data hasil. Ini memungkinkan pandangan holistik tentang pilihan intervensi, termasuk dampak keuangan, lingkungan dan sosial yang luas.	Dapat menangkap dampak jangka pendek dan jangka panjang secara sistematis. Kualitas hasil sangat bergantung pada data yang tersedia dan kemampuan untuk memonetisasi dampak.

Tabel 5. Panduan cepat untuk metode evaluasi *value-for-money*

3.6. Metode Sintesis

Metode sintesis memerlukan penggabungan temuan dari berbagai studi ke dalam pemahaman umum tentang dampak dan/atau penyampaian suatu kebijakan. Sintesis pasca evaluasi digunakan untuk menyatukan keluaran dari berbagai metode yang digunakan dalam evaluasi dengan cara yang jelas dan sistematis untuk menjawab pertanyaan evaluasi awal. Metode komponen dapat berupa metode kuantitatif atau kualitatif atau metode campuran dan berfokus pada pertanyaan evaluasi dampak atau proses.

Untuk mensintesis, temuan evaluasi diintegrasikan di sekitar setiap pertanyaan (bukan hanya melaporkan temuan dari setiap metode), biasanya disebut 'triangulasi'. Idealnya, ini memberikan konsensus bukti dengan kepastian yang lebih besar daripada masing-masing bagian komponen. Namun, jika ada bukti yang bertentangan, temuan harus diperiksa dengan hati-hati, penjelasan dipertimbangkan, dan bukti tambahan dicari jika sesuai.

Metode sintesis juga dapat digunakan sebelum memulai pengembangan intervensi untuk menentukan apa yang sudah diketahui pada suatu subjek.

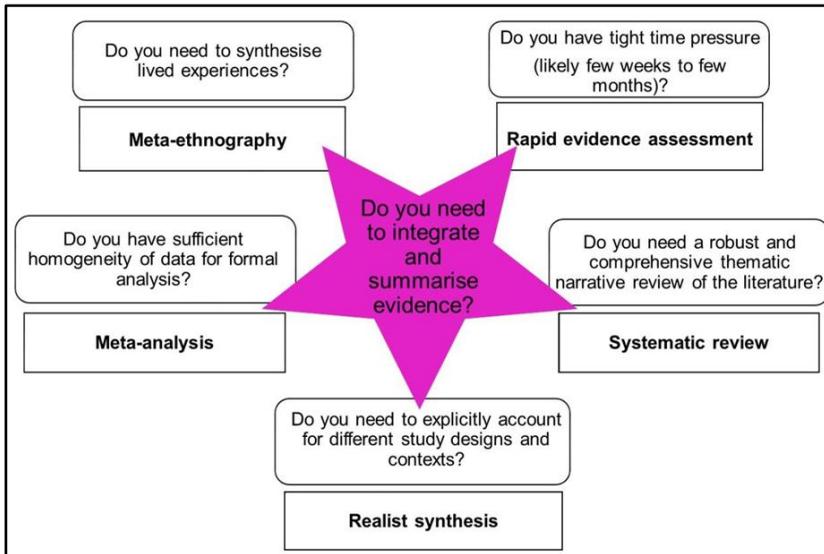
Meta-evaluasi dan sintesis adalah metode sintesis formal yang dirancang untuk menyatukan temuan dari sejumlah studi berbeda tentang subjek yang sama. Metode sintesis dan meta-analisis menyatukan bukti evaluasi yang ada secara terstruktur untuk menghasilkan narasi yang koheren. Metode 'penelitian sekunder' ini cenderung dimulai dengan serangkaian tujuan dan kriteria inklusi/eksklusi yang dinyatakan dengan jelas. Mereka kemudian menggunakan protokol eksplisit dan kriteria yang telah ditentukan sebelumnya untuk menilai kualitas studi sumber, mengekstrak temuan kunci dari studi tersebut, dan menganalisis hasil untuk memberikan akun terintegrasi dari literatur sumber.

Keterbatasan utama sintesis evaluasi adalah ketergantungannya pada kumpulan data yang diperoleh sebelumnya. Kualitas sintesis bukti tergantung pada kualitas studi asli, dan validitas tergantung pada studi asli mana yang disertakan. Oleh karena itu, pentingnya kriteria inklusi dan penilaian kualitas dalam prosesnya.

Sebagai contoh, Tinjauan Sistematis dan Penilaian Bukti Cepat mengintegrasikan bukti literatur untuk menghasilkan ringkasan naratif. Tinjauan sistematis sering memakan waktu berbulan-bulan dan oleh karena itu, Penilaian Bukti Cepat dapat

digunakan jika ketelitian yang lebih rendah dapat ditoleransi. Metode ini sering berfokus pada data kuantitatif, tetapi kerangka kerja lainnya dapat digunakan untuk mengintegrasikan informasi kualitatif secara sistematis.

Serangkaian pertanyaan diberikan di bawah ini pada Gambar 3 untuk membantu mempertimbangkan metode sintesis bukti mana yang paling tepat.



Gambar 3. Memilih metode sintesis bukti

BAB 4.

PENGUMPULAN, SUMBER DAN KETERKAITAN DATA

Pengumpulan data merupakan komponen penting dari setiap evaluasi dan perlu direncanakan terlebih dahulu. Merencanakan pengumpulan data bersamaan dengan pengembangan intervensi memastikan bahwa pengumpulan data dan akses data dibangun ke dalam desain kebijakan dan undang-undang terkait.

Penting untuk dicatat bahwa data dasar – data yang dikumpulkan sebelum intervensi – perlu dikumpulkan lebih awal; dan bahwa data perbandingan – data dari kelompok yang tidak terpengaruh oleh intervensi – perlu dinegosiasikan. Tanpa perencanaan pengumpulan data atau akses data yang tepat, evaluasi mungkin tidak mungkin dilakukan, sangat terbatas, atau tidak perlu mahal. Jika pengumpulan data dirancang dengan buruk, ini dapat mengakibatkan pengumpulan data yang tidak akurat dan kesimpulan yang salah diambil dari evaluasi.

Teori Perubahan dapat digunakan untuk mengidentifikasi kebutuhan dan kesenjangan data. Kombinasi pengumpulan data yang ada dan data baru sering kali dapat memberikan pemahaman dan penelusuran yang paling lengkap tentang Teori Perubahan dan efektivitas intervensi. Pengumpulan data menggunakan beberapa metode, seperti data pemantauan dan survei yang dipesan lebih dahulu, membangun kepercayaan pada temuan dan ketahanan terhadap evaluasi; ini dikenal sebagai triangulasi.

Dalam banyak kasus, data berhubungan dengan individu tetapi dapat berhubungan dengan unit lain, misalnya sekolah, bisnis, dan wilayah geografis. Pertimbangan serupa berlaku dalam semua kasus.

4.1. Penentuan Jenis Data

Berikut ini terdapat beberapa pertanyaan kunci, serta pertimbangan tentang pengumpulan data dalam merencanakan evaluasi.

Pertanyaan 1.

Jenis data apa yang diperlukan untuk menjawab setiap pertanyaan evaluasi?

Pertimbangan:

Apa yang diperlukan untuk menjawab pertanyaan evaluasi yang diidentifikasi selama tahap pelingkupan. Jenis data meliputi:

- Data numerik

- Data dokumenter (data atau informasi yang sudah dikumpulkan)
- Data observasi
- Deskripsi pengalaman, pendapat, dan pandangan masyarakat
- Kombinasi/triangulasi bukti di atas.

Pertanyaan 2.

Siapa atau apa yang dapat memberikan data ini?

Pertimbangan:

- Peserta intervensi
- Penyedia jasa
- Pemangku Kepentingan
- Database
- Survei yang ada
- Survei dipesan lebih dahulu

Pertanyaan 3.

Apakah ada masalah dengan mengakses atau mengumpulkan data?

Pertimbangan:

- Masalah akses data (misalnya masalah hukum, prosedur internal, identifikasi kelompok sasaran, mengumpulkan data pembanding)
- Sensitivitas data/masalah etika (misalnya meneliti sensitif populasi, masalah akses data)
- Ketersediaan kerangka sampling yang diperlukan
- Potensi 'beban data' pada responden: apakah proporsional meminta orang untuk memberikan data?
- Siapa yang bertanggung jawab atas pengumpulan data?
- Apakah tugas pengumpulan data sebanding dengan nilainya? Apakah metode lain mungkin?
- Apakah tindakan pengumpulan data dari individu dapat memengaruhi perilaku?

Pertanyaan 4.

Dari bagian mana dari populasi yang menarik data harus dikumpulkan?

Pertimbangan:

Ada beberapa cara untuk mengumpulkan data dari populasi yang diminati. Penting untuk mempertimbangkan apa yang proporsional untuk menjawab pertanyaan evaluasi. Pilihan meliputi:

- Sensus dari semua data/populasi minat yang tersedia

- Sampel representatif dari data yang tersedia/populasi
- Subset populasi yang diminati, dipilih secara sengaja untuk mencakup berbagai konteks tetapi tidak representatif secara statistik.

Pertanyaan 5.

Bagaimana metode analisis data yang diusulkan akan mempengaruhi data yang dibutuhkan?

Pertimbangan:

Jenis analisis yang akan diterapkan pada data biasanya akan menentukan sifat dan kuantitas data yang dibutuhkan. Sebagai contoh:

- Tingkat presisi yang diperlukan
- Sub-populasi yang diinginkan (dan oleh karena itu, ukuran sampel yg dibutuhkan)
- Sifat metode yang akan digunakan dan oleh karena itu, persyaratan data
- Apakah data dasar atau data kelompok kontrol diperlukan.

4.2 Sumber Data

Ada banyak sumber yang berbeda untuk dipertimbangkan ketika mencari untuk mengumpulkan data untuk evaluasi. Ini termasuk:

- a) Data administratif dan pemantauan yang ada: data yang dikumpulkan baik untuk operasi intervensi, atau untuk fungsi lainnya.
- b) Data survei skala besar yang ada: data survei skala besar jangka panjang, sering dikelola oleh pemerintah pusat, Kantor Statistik Nasional atau badan pendanaan penelitian, seperti UKRI (misalnya data dari Survei Angkatan Kerja atau Survei Kejahatan untuk Inggris dan Wales).
- c) Sumber data baru yang dirancang khusus untuk evaluasi: data tersebut biasanya diperoleh melalui metode seperti survei, metode kualitatif (wawancara, observasi, kelompok fokus), web-scraping.
- d) Data media sosial: ini adalah sumber data yang relatif baru, tetapi berpotensi sangat kaya dalam hal mengukur reaksi yang tidak diinginkan terhadap suatu intervensi.

Jika memungkinkan, data administratif atau data lain yang ada harus digunakan karena ini akan hemat biaya dan seringkali mencakup seluruh populasi yang diminati. Namun, data baru seringkali diperlukan untuk menjawab pertanyaan spesifik, khususnya yang tidak penting untuk manajemen kinerja suatu program.

Berbagai macam metode pengumpulan data kuantitatif dan kualitatif dapat digunakan. Masing-masing metode pengumpulan data sekarang akan dijelaskan secara bergiliran meliputi tujuan, penerapan dan pertimbangannya dalam suatu evaluasi.

4.2.1. Data administrasi/ monitoring

Data administratif dikumpulkan, disimpan, dan digunakan terutama untuk tujuan administratif (non-penelitian). Data administratif biasanya dikumpulkan pada pendaftaran, transaksi, dan penyimpanan catatan, biasanya selama pemberian layanan (misalnya: Basis Data Siswa Nasional yang dipegang oleh DfE; Komputer Nasional Polisi yang dipegang oleh Kementerian Kesehatan; dan catatan dari penyedia layanan tentang pelanggan mereka dan pelayanan yang disediakan). Kumpulan data administratif biasanya digunakan untuk membentuk dasar publikasi Statistik Resmi.

Ketersediaan data administratif atau data dari survei jangka panjang yang ada harus selalu dipertimbangkan pada tahap desain evaluasi. Kedua jenis data tersebut dapat berfungsi sebagai sumber penting dari latar belakang atau data penjelas yang mendahului intervensi yang memberikan informasi tren yang sangat baik. Jika memungkinkan, penggunaan data yang ada mengurangi biaya keuangan evaluasi dan beban responden.

Namun, kehati-hatian perlu dilakukan saat menggunakan data ini untuk tujuan evaluasi karena data mungkin tidak fokus pada pertanyaan spesifik yang diminati atau dikumpulkan secara teratur untuk memberikan bukti yang berguna terhadap pertanyaan evaluasi.

Dalam beberapa kasus, data administratif mengukur konsep yang terkait, tetapi tidak identik, dengan ukuran kepentingan. Misalnya, meskipun data administratif dapat memberi tahu kita dengan sangat tepat dan akurat siapa yang menerima tunjangan, mengukur definisi 'pengangguran' yang paling diterima secara luas memerlukan data survei. Meskipun kedua konsep ini terkait erat, mereka tidak sama.

Data pemantauan atau data manajemen kinerja dikumpulkan selama intervensi untuk memberikan jawaban atas sejumlah pertanyaan kebijakan, penelitian, dan kinerja.

Data pemantauan biasanya mencakup semua aspek operasi intervensi; misalnya, informasi tentang mereka yang mengakses layanan, serta masukan, proses, keluaran dan hasil.

Data pemantauan seringkali bersifat administratif dan kuantitatif, dan

seringkali tidak dibuat terutama untuk tujuan evaluasi. Biasanya digunakan untuk membantu melacak kemajuan penyampaian intervensi; atau untuk mengidentifikasi di mana intervensi tidak dilaksanakan seperti yang diharapkan dan tindakan lebih lanjut apa yang diperlukan untuk memastikannya dapat mencapai tujuannya. Data pemantauan dapat, jika ditentukan dan dikumpulkan dengan tepat, digunakan untuk mengukur manfaat sebagai bagian dari proses manajemen manfaat formal.

4.2.1. Data survei

Data survei skala besar sering dikumpulkan oleh departemen dan penelitian pemerintah lembaga untuk tujuan statistik dan penelitian.

Data survei skala besar biasanya memiliki sampel representatif yang besar, yang memberikan perkiraan yang kuat. Menggunakan survei ini bisa lebih murah daripada mengumpulkan data baru. Namun, ada potensi kerugian:

- tujuan pengumpulan data akan berbeda dengan tujuan evaluasi;
- mereka tidak mungkin memberikan rincian yang dibutuhkan; yaitu, jawaban untuk spesifik pertanyaan atau informasi tentang populasi yang diinginkan (misalnya, total sampel survei mungkin besar, tetapi jumlah orang dalam sampel yang ditargetkan oleh intervensi mungkin kecil dan tidak mungkin untuk mengidentifikasi mereka secara akurat);
- waktu survei akan ditentukan dan mungkin tidak sesuai dengan evaluasi tujuan.

Namun, mereka seringkali merupakan sumber latar belakang atau data eksplorasi yang sangat baik; misalnya, memberikan prevalensi kegiatan/perilaku/sikap tertentu pada populasi umum, yang dapat berguna dalam membongkar masalah atau menjelaskan konteksnya.

Pengumpulan data kuantitatif baru biasanya melibatkan survei sampel populasi untuk membuat perkiraan tentang populasi yang lebih luas. Mungkin tepat untuk melakukan sensus terhadap seluruh penduduk; atau dalam beberapa kasus, menanggapi dapat dibuat wajib sebagai bagian dari syarat untuk mengambil bagian dalam suatu intervensi. Dalam hal ini, kemungkinan bias non-respons akan lebih kecil.

Contoh

Metode pengambilan sampel terbagi dalam dua kategori besar: sampling probabilitas dan sampling nonprobabilitas.

Sebagian besar survei yang digunakan dalam evaluasi akan didasarkan pada metode pengambilan sampel probabilitas, yang melibatkan pemilihan responden secara acak dari kerangka sampel (yaitu daftar semua responden dalam populasi yang diinginkan).

Metode utama untuk mendapatkan sampel probabilitas adalah:

- Pengambilan sampel acak sederhana: Melibatkan penggunaan metode acak, biasanya dihasilkan komputer, untuk memilih individu, dengan semua individu memiliki probabilitas yang sama untuk seleksi.
- Pengambilan sampel bertingkat: Membagi populasi ke dalam kelompok-kelompok kemudian memilih sampel probabilitas dari setiap kelompok, misalnya berdasarkan lokasi geografis. Probabilitas pemilihan dapat bervariasi antar kelompok untuk memastikan bahwa ada sampel yang cukup besar dari masing-masing kelompok.
- Pengambilan sampel cluster: Cluster dipilih menggunakan metode probabilitas dan hanya individu dalam cluster yang dipilih. Hal ini dapat membantu untuk mengurangi biaya kerja lapangan, terutama dalam survei tatap muka di mana cluster biasanya wilayah geografis tetapi meningkatkan ketidakpastian perkiraan yang berasal dari data.
- Pengambilan sampel multi-tahap: Melibatkan penggunaan kombinasi metode pengambilan sampel.

Asalkan tidak ada masalah dengan non-respons (lebih lanjut tentang ini di akhir bagian ini), jenis pengambilan sampel ini akan memastikan bahwa perkiraan yang diperoleh dari survei mewakili populasi secara keseluruhan (atau dapat diberi bobot untuk membuatnya jadi), dan kesalahan pengambilan sampel dari perkiraan tersebut dapat dihitung.

Pendekatan non-probabilitas dapat nyaman dan lebih murah tetapi harus digunakan dengan sangat hati-hati karena tidak mungkin untuk menyimpulkan dengan yakin apa pun tentang populasi secara keseluruhan dari hasil survei.

Pengambilan sampel kuota adalah pendekatan non-probabilitas yang banyak digunakan, yang melibatkan pengaturan 'kuota' untuk berbagai jenis responden. Hal ini memungkinkan kontrol yang lebih besar atas komposisi sampel, seperti memastikan representasi yang memadai dari kelompok yang berbeda, atau pengambilan sampel yang berlebihan dari kelompok minoritas

yang tidak perlu diidentifikasi dalam kerangka pengambilan sampel. Sampel kuota yang dirancang dengan baik adalah alternatif yang valid untuk pengambilan sampel probabilitas. Desain sampel kuota terkuat dimulai dengan sampel yang diambil secara acak dari kerangka sampling yang representatif dan kemudian menetapkan kuota untuk jumlah wawancara yang ingin dicapai dengan kelompok tertentu. Itu harus memiliki prosedur kerja lapangan yang mengurangi risiko hanya termasuk responden yang paling mudah untuk mendapatkan tanggapan dari. Desain yang lebih lemah dapat melibatkan 'contoh praktis' dari individu yang mudah dihubungi, atau 'sampel bola salju'. Panduan pemerintah lebih lanjut tentang pengambilan sampel kuota tersedia²².

Semua desain survei perlu mempertimbangkan dengan cermat bagaimana memaksimalkan tingkat

respons dan meminimalkan bias non-respons, seperti melalui protokol kerja lapangan yang menentukan jumlah upaya yang harus dilakukan untuk menghubungi orang. Jika tingkat non-respons tidak terdistribusi secara merata dalam sampel probabilitas, selain yang terkait dengan karakteristik yang diamati dalam kerangka sampling, ini dapat menyebabkan hasil yang bias, dan kepercayaan pada kesimpulan tentang populasi yang lebih luas akan berkurang.

4.3.3. Data Kualitatif

Metode pengumpulan data kualitatif memberikan pemahaman mendalam tentang perilaku, persepsi dan alasan yang mendasari fenomena sosial. Sementara metode kuantitatif biasanya digunakan untuk mengukur 'apa', metode kualitatif paling sering digunakan untuk mengeksplorasi 'bagaimana' dan 'mengapa'.

Metode pengumpulan data kualitatif yang umum meliputi:

- Wawancara mendalam: Ini digunakan untuk mengumpulkan data tentang sejarah pribadi, perspektif, dan pengalaman, terutama ketika topik sensitif sedang dieksplorasi, atau masalah yang sedang dibahas tidak dipahami dengan baik. Pertanyaan cenderung 'terbuka' yang memungkinkan tanggapan terperinci. Wawancara mendalam biasanya dilakukan secara tatap muka atau melalui telepon.
- Kelompok terarah: Peserta kelompok terarah didorong untuk berdiskusi dan berdebat secara terbuka untuk memahami pandangan dan pengalaman, memungkinkan berbagai pandangan untuk dieksplorasi. Ini dapat digunakan

secara efektif untuk perencanaan tindakan dan mengembangkan atau meningkatkan produk dan layanan.

- Studi kasus: Ini adalah investigasi mendalam, mungkin jangka panjang, dari satu masalah atau sejumlah kecil orang, peristiwa, konteks, bidang, organisasi atau kebijakan.
- Observasi: Suatu proses mengamati subjek penelitian (dengan persetujuan dan pengetahuan mereka) untuk mengamati perilaku mereka tanpa menyanjai mereka. Dapat digunakan secara efektif ketika menguji coba prosedur atau proses baru untuk mengukur tindakan dan perilaku responden.
- Etnografi: Ini termasuk observasi, tetapi juga partisipasi. Etnografi berusaha untuk memahami orang dan bagaimana mereka hidup dalam lingkungan budaya dan fisik mereka. Wawancara etnografi akan berbeda dari wawancara mendalam yang lebih tradisional karena peneliti biasanya akan berbagi waktu dan membangun hubungan dan kepercayaan dengan orang yang diwawancarai. Etnografi film adalah salah satu contoh bagaimana pendekatan ini digunakan oleh Lab Kebijakan dan Pemerintah²⁸.

Memilih responden

Untuk pengumpulan data kualitatif, pemilihan responden bertujuan untuk memastikan data menangkap kekayaan pandangan dan pendapat dalam populasi yang diminati; atau oleh menggambarkan melalui eksplorasi yang sangat mendalam cerita atau konteks tertentu. Penting untuk mempertimbangkan sudut pandang mana yang harus dimasukkan dalam evaluasi karena ini akan berdampak pada ukuran sampel.

Tidak ada aturan praktis yang disepakati untuk berapa banyak responden yang harus dipilih. Tujuannya bukan untuk dapat menghasilkan temuan yang secara statistik mewakili populasi, tetapi untuk memahami berbagai pandangan tentang topik tertentu, dan untuk memberikan wawasan yang lebih dalam tentang fenomena sosial. Mencapai titik jenuh - titik di mana penambahan data baru tidak menambah sesuatu yang baru pada temuan merupakan indikasi yang baik bahwa luasnya pandangan dan pendapat telah ditangkap dan wawancara tambahan tidak mungkin menambah nilai.

4.2.4. Data media sosial

Penggunaan media sosial sebagai sumber data untuk evaluasi semakin meningkat dalam beberapa tahun terakhir, baik untuk analisis kualitatif maupun kuantitatif. Bagaimana media sosial digunakan sebagai sumber data ditentukan oleh jenis perilaku yang diteliti dan platform yang digunakan sebagai sumber materi. Pengembangan alat otomatis (pengikisan web) memungkinkan sejumlah besar data dikumpulkan, dibersihkan, disimpan, dan dianalisis dengan cepat. Penggunaan data media sosial menghadirkan tantangan khusus sebagai sampel, meskipun sangat besar, dipilih sendiri, dan tidak selalu mudah untuk menilai keandalan kemungkinan akurasi data.

4.3 Kualitas Data

Kualitas data sangat penting untuk menghasilkan temuan kuat yang dapat dengan jelas dan diinterpretasikan secara tepat oleh para pengambil keputusan. Kualitas data yang ada harus dinilai sejak awal; kualitas data yang buruk atau sebagian akan mempengaruhi cakupan dan skala kontribusi data terhadap suatu evaluasi. Transparansi adalah kuncinya dan pemeriksaan kualitas harus dimasukkan ke dalam desain evaluasi.

Kualitas data akan tergantung pada jenis pertanyaan yang diajukan dan jenis alat pengumpulan data yang digunakan.

Data tentang masalah yang berpotensi sensitif, seperti orientasi seksual atau disabilitas, di mana orang yang berbeda dapat menafsirkan pertanyaan dengan cara yang berbeda, seringkali lebih dapat diandalkan jika dikumpulkan sebagai bagian dari latihan penelitian daripada melalui sistem pemantauan.

Di sisi lain, data administratif atau pemantauan akan memiliki kualitas yang lebih tinggi saat memeriksa masalah seperti tanggal pasti bergabung atau keluar dari suatu program. Data tersebut adalah lebih mungkin untuk direkam secara akurat pada sistem pemantauan daripada diingat oleh peserta dalam sebuah wawancara. Dalam beberapa kasus, biasanya ketika digunakan secara langsung sebagai bagian dari sistem pembayaran, data tunduk pada audit formal dan oleh karena itu, sangat dapat diandalkan.

4.4. Pengolahan Data

Semua data harus dikumpulkan, ditangani, dan diakses dengan tepat.

Setiap orang yang menangani data harus dilatih untuk memahami bagaimana melakukan ini. Semua data yang digunakan dalam evaluasi, terlepas dari sumbernya, harus dikumpulkan, ditransfer, disimpan, diproses, dan dihapus sesuai dengan GDPR, dan proses keamanan departemen tertentu. Ini berlaku untuk organisasi eksternal, seperti kontraktor penelitian, yang mengumpulkan data atas nama departemen. Protokol akses data harus mencakup isu-isu, seperti: bagaimana mengotorisasi akses ke data oleh kelompok yang berbeda; akses jarak jauh; akses Wi-Fi; pelatihan penanganan dan keamanan data; dan menutupi/menkripsi data.

Data dapat disimpan dalam berbagai format, seperti dalam database, dataset, spreadsheet, atau gudang data. Memastikan akses data yang tepat – mereka yang dapat membaca, menulis, menyimpan, dan mengubah data – sangat penting dan harus dibatasi untuk mereka yang memiliki hak hukum dan persyaratan kerja.

Risiko penanganan data serius dan mencakup: pelanggaran data kriminal; individu atau kelompok yang menderita kerugian atau rasa malu yang parah; gangguan yang tidak dapat dibenarkan; hilangnya privasi; kehilangan atau kompromi data; tantangan hukum; dan kerusakan reputasi di tingkat departemen, Pegawai Negeri Sipil dan Pemerintah yang lebih luas.

Data harus dianonimkan pada kesempatan pertama – yaitu, pengenalan langsung, seperti nama dan alamat, harus dihapus dari file analitik. Jika perlu untuk menyimpan ini, setidaknya untuk sementara, mereka harus disimpan secara terpisah dengan akses yang dikontrol secara ketat. Meski begitu, dalam banyak kasus data tersebut kemungkinan akan tetap menjadi data pribadi dan harus dilindungi dengan hati-hati.

Di akhir proyek, jika memungkinkan, data harus sepenuhnya dianonimkan; sebuah proses yang kemungkinan akan mencakup, misalnya, mengganti tanggal lahir yang tepat dengan bulan dan tahun lahir, atau kode pos yang tepat dengan pengkodean geografis yang lebih luas. Jika hal ini tidak memungkinkan, dan dianggap perlu atau diinginkan untuk menyimpan data pribadi di luar umur proyek, pengamanan yang sesuai harus diambil.

BAB 5.

DISEMINASI HASIL EVALUASI

Nilai evaluasi datang melalui penggunaan dan pengaruhnya. Ini harus dipertimbangkan pada tahap perencanaan dan terus ditinjau kembali selama pelaksanaan evaluasi.

Pengguna evaluasi bisa langsung atau tidak langsung.

- **Langsung:** mereka yang merancang dan mengimplementasikan intervensi menggunakan temuan untuk meningkatkan desain atau implementasi intervensi dan untuk memaksimalkan peluang mencapai hasil yang diinginkan. Mereka yang meneliti kinerja pemerintah menggunakan temuan evaluasi untuk menilai kinerja dan hasil kebijakan.

- **Tidak langsung:** mereka yang menggunakan temuan untuk menjawab pertanyaan lain yang terkait, merancang kebijakan masa depan di bidang yang sama atau serupa, atau menangkap pembelajaran tentang penggunaan dana publik.

Keluaran evaluasi perlu dirancang dengan hati-hati untuk memenuhi kebutuhan berbagai pengguna. Evaluasi dengan implikasi yang jelas dan terbukti untuk pengambilan keputusan di masa depan hanya bernilai jika temuan dapat diakses dan digunakan oleh pembuat keputusan yang relevan pada waktu yang tepat. Demikian pula, badan pengawasan dan masyarakat luas harus dapat dengan mudah mengakses, memahami, dan menggunakan temuan untuk menilai desain, implementasi, dan hasil kebijakan pemerintah. Tantangan untuk mencapai penggunaan dan pengaruh adalah umum bagi semua organisasi yang ingin mengekstrak nilai maksimum dari penelitian mereka.

5.1. Pengembangan Rencana Diseminasi

Tidak ada cara pasti untuk mengembangkan penggunaan evaluasi dan rencana diseminasi. Kuncinya adalah mempertimbangkan empat pertanyaan di bawah ini:

- Kelompok yang mana? Daftar pemangku kepentingan potensial dapat ditemukan di bagian 1.10. Penting untuk mempertimbangkan semua pengguna potensial secara individual karena mudah untuk melewatkan pemangku kepentingan utama yang mungkin memiliki kebutuhan yang sangat berbeda dari evaluasi. Latihan pemetaan pemangku kepentingan dapat bermanfaat.
- Informasi apa? Mengetahui informasi apa yang dibutuhkan oleh masing-masing kelompok harus menginformasikan pertanyaan evaluasi yang diajukan.
- Titik waktu yang mana? Akan ada poin keputusan selama implementasi dan penyampaian intervensi apa pun. Mengetahui poin-poin ini dapat memungkinkan berbagai rencana pengumpulan data dirancang untuk memenuhi persyaratan ini. Catatan: kebutuhan bukti sering kali muncul sebelum titik keputusan akhir untuk menginformasikan pemikiran.
- Untuk tujuan apa? Bukti evaluasi dapat digunakan untuk berbagai tujuan, yang akan berdampak pada jenis bukti yang dihasilkan dan bagaimana bukti tersebut digunakan dan disebarluaskan. Ini sangat penting untuk poin keputusan sementara. Ketika keputusannya kecil (misalnya, tinjauan jumlah staf untuk mendukung intervensi), data pemantauan yang muncul mungkin sudah cukup. Namun, jika keputusannya besar (misalnya, apakah akan menggelar intervensi secara nasional), kebutuhan bukti akan jauh lebih besar dan memerlukan tingkat ketahanan yang jauh lebih tinggi.

Mengidentifikasi persyaratan ini pada tahap desain evaluasi memungkinkan diskusi, negosiasi tentang penjadwalan dan pengelolaan ekspektasi sejak awal.

Rencana tersebut harus mencakup apa yang akan diterbitkan, kapan akan diterbitkan dan alat komunikasi apa yang akan digunakan (misalnya publikasi cetak, publikasi keras, media sosial, atau seminar dan presentasi konferensi). Ini harus mencakup poin keputusan dan bukti apa yang akan tersedia kapan. Semua keluaran yang akan dipublikasikan harus tercakup, termasuk laporan, data dasar dan alat penelitian.

Merinci 'tujuan pengaruh' dalam rencana penggunaan dan diseminasi dapat bermanfaat. Pengaruh tujuan menentukan dampak yang diinginkan dari

evaluasi. Misalnya, evaluasi dapat dilakukan untuk mengidentifikasi praktik yang baik dalam skema rehabilitasi penjara. Tujuan pengaruh dapat berupa penggunaan temuan evaluasi untuk secara positif mengubah cara pelaksanaan program rehabilitasi. Sekali mempengaruhi tujuan yang jelas, hasil evaluasi dapat disesuaikan untuk memenuhi tujuan ini.

Untuk mendapatkan dukungan, adalah praktik yang baik untuk menyetujui rencana penggunaan dan diseminasi dengan sebanyak mungkin pemangku kepentingan; ini biasanya dicakup oleh papan program dan kelompok pengarah. Keterlibatan awal ini juga dapat membantu memprioritaskan dan mengelola harapan karena tidak mungkin semua pertanyaan yang diajukan dapat dijawab.

Dua kelompok tambahan yang berguna untuk didekati ketika mengembangkan rencana penggunaan dan penyebaran adalah kantor pers departemen (untuk menyetujui prinsip-prinsip umum dalam rencana dan pendekatan luas yang akan digunakan) dan kantor menteri (untuk mendapatkan dukungan dan persetujuan mereka untuk rencana publikasi di muka). Ini adalah praktik yang baik untuk menghindari persepsi bahwa keputusan publikasi terlalu dipengaruhi oleh sifat temuan.

Setelah audiens yang berbeda dan kebutuhan bukti mereka diketahui dan diprioritaskan, pelaporan dan komunikasi harus disesuaikan untuk memenuhi kebutuhan ini. Nilai dari menyesuaikan komunikasi dengan cara ini tidak dapat diremehkan. Meningkatkan kegunaan temuan dengan membantu audiens tertentu memahami bagaimana temuan berhubungan langsung dengan bidang minat mereka dapat sangat berharga dalam memastikan temuan digunakan.

Meningkatkan penggunaan temuan melibatkan pertimbangan kelompok mana yang memiliki kapasitas untuk bertindak berdasarkan temuan, dengan cara apa dan kendala yang mungkin mereka hadapi.

5.2. Diseminasi Hasil Evaluasi

Penting untuk mengidentifikasi dan menggunakan keluaran dan saluran komunikasi yang paling tepat untuk memaksimalkan jangkauan, dan berdampak pada, berbagai kelompok pemangku kepentingan evaluasi.

Sementara laporan evaluasi akhir merupakan dokumen penting dalam studi evaluasi, jarang merupakan cara terbaik untuk memastikan dampak dan pengaruh. Misalnya, staf dan menteri pengiriman garis depan akan memiliki kebutuhan informasi dan gaya komunikasi yang sangat berbeda. Mampu

memasukkan temuan yang muncul ke dalam diskusi dewan program dapat sangat berharga bagi pertimbangan kelompok.

Mode komunikasi alternatif termasuk ringkasan satu halaman, output video, infografis, berbagi data, buletin, posting media sosial, presentasi konferensi, dan seminar. Membahas kebutuhan bukti dengan pemangku kepentingan dapat membantu mengidentifikasi saluran komunikasi yang paling efektif untuk kelompok tersebut. Perlu dicatat bahwa tidak tepat untuk membahas temuan di depan umum sebelum dipublikasikan dalam beberapa bentuk. Namun, memastikan pemangku kepentingan, seperti kelompok pengarah dan pemimpin kebijakan/penyampaian, memahami pesan sebelum publikasi dapat bermanfaat, terutama dalam penanganan temuan negatif.

Publikasi dapat menjadi sumber daya yang intensif bagi manajer evaluasi pemerintah karena memerlukan jaminan kualitas, pengeditan, izin menteri dan masukan dari kantor pers.

Jika sesuai, opsi dapat berupa menerbitkan keluaran reguler, pendek, untuk memastikan temuan dirilis dan dapat digunakan secara terbuka pada waktu yang tepat.

Kontraktor akademik dapat berusaha untuk mempublikasikan hasil evaluasi dalam jurnal akademik, yang memiliki keuntungan mengekspos pekerjaan ke lebih banyak orang, memberikan jaminan kualitas lebih lanjut dan penyebaran pengetahuan. Mendorong hal ini pada awal proyek juga berpotensi menarik lebih banyak akademisi untuk berpartisipasi dalam kegiatan evaluasi.

5.3. Membangun Budaya Evaluasi

Memahami evaluasi sebagai praktik yang melekat berarti bahwa evaluasi tidak boleh beroperasi sebagai fungsi yang terpisah dan independen, tetapi lebih sebagai bagian yang terintegrasi dari budaya organisasi dan struktur operasional. Ini berarti bahwa praktik evaluatif dan reflektif adalah bagian dari 'cara kita melakukan sesuatu di sekitar sini' di mana semua rekan kerja mencari, belajar, dan berpikir kritis tentang bukti yang mendukung tindakan mereka.

Menanamkan cara berpikir ini jarang terjadi secara kebetulan. Ini menyiratkan pergeseran dari melihat evaluasi sebagai kejahatan yang diperlukan atau kebutuhan ritualistik, ke evaluasi di mana menghubungkan nilai dengan implementasi kebijakan dibangun sebagai dimensi inti dari praktik kebijakan yang baik. Ini sangat menyiratkan 'menggunakan' evaluasi untuk memandu tindakan di masa depan dan 'langkah selanjutnya' dan sangat relevan untuk bekerja di lingkungan yang kompleks.

Dalam istilah sehari-hari, itu berarti kepala departemen atau inisiatif khusus perlu menekankan harapan bahwa dimensi evaluatif diperlukan pada tahap desain kebijakan. Kolega harus didorong secara aktif untuk mempertimbangkan bagaimana nilai suatu kebijakan akan ditetapkan sebagai hal yang biasa.

BAB 6.

EVALUASI PROGRAM KESEHATAN MASYARAKAT

6.1. Alasan Perlunya Mengevaluasi Program Kesehatan Masyarakat

Petugas evaluasi mungkin mengevaluasi untuk melakukan evaluasi oleh mandat eksternal dari orang dana, pemberi kuasa, atau orang lain, atau mereka mungkin ditarik untuk melakukan evaluasi oleh kebutuhan internal untuk menentukan bagaimana kinerja program dan apa yang dapat ditingkatkan. Sementara dorongan atau tarikan dapat memotivasi sebuah program untuk melakukan yang baik, upaya evaluasi program lebih mungkin dipertahankan ketika staf melihat hasilnya sebagai informasi yang berguna yang dapat membantu mereka melakukan pekerjaan mereka dengan lebih baik.

Data yang dikumpulkan selama evaluasi memungkinkan manajer dan staf untuk membuat program sebaik mungkin, belajar dari kesalahan, membuat modifikasi yang diperlukan, menyatukan

kemajuan menuju tujuan program, dan menilai keberhasilan program dalam mencapai jangka pendek, menengah, dan hasil jangka panjang. sebagian besar program kesehatan masyarakat bertujuan untuk mengubah perilaku dalam satu atau lebih kelompok

dan untuk menciptakan lingkungan yang memperkuat adopsi berkelanjutan dari perubahan ini, dengan maksud bahwa perubahan lingkungan dan perilaku akan mencegah dan mengendalikan penyakit dan cedera.

Melalui evaluasi, Anda dapat melacak perubahan ini dan, dengan desain evaluasi yang cermat, menilai efektivitas dan dampak program, atau strategi tertentu dalam menghasilkan perubahan ini.

Rekomendasi yang relevan dengan evaluasi program kesehatan masyarakat yang komprehensif meliputi:

- Mendorong pendekatan evaluasi partisipatif yang memberikan kesempatan yang melibatkan keterlibatan oleh semua orang yang memiliki kepentingan langsung dalam inisiatif (program, kebijakan, dan kegiatan terorganisir lainnya).
- Mengharuskan sebagian dari total sumber daya keuangan untuk inisiatif promosi kesehatan dialokasikan untuk evaluasi—mereka merekomendasikan 10%.
- Memastikan bahwa campuran informasi proses dan hasil yang digunakan untuk semua inisiatif promosi kesehatan.

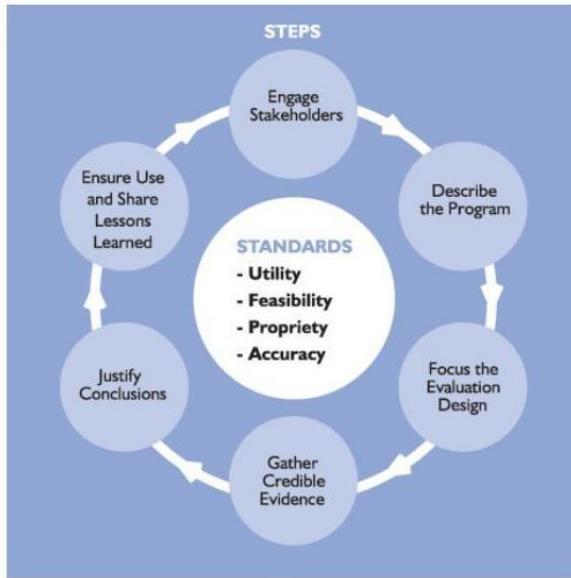
- Mendukung penggunaan berbagai metode untuk memulai inisiatif promosi kesehatan.
- Mendukung penelitian lebih lanjut ke dalam pengembangan pendekatan yang tepat untuk memulai promosi kesehatan.
- Mendukung pengembangan pelatihan dan pendidikan untuk mengembangkan keahlian dalam evaluasi inisiatif promosi kesehatan.
- Menciptakan dan mendukung kesempatan untuk berbagi informasi tentang metode evaluasi yang digunakan dalam Melalui konferensi, lokakarya, promosi jaringan, dan sarana lainnya.

6.2. Kerangka Evaluasi Program dalam Kesehatan Masyarakat

Evaluasi program adalah salah satu dari sepuluh layanan kesehatan masyarakat yang penting dan praktik organisasi yang penting dalam kesehatan masyarakat. Sampai saat ini, bagaimanapun, hanya ada sedikit kesepakatan di antara pejabat kesehatan masyarakat tentang prinsip dan prosedur untuk melakukan studi semacam itu.

Pada tahun 1999, CDC menerbitkan kerangka untuk Evaluasi Program dalam Kesehatan Masyarakat dan beberapa rekomendasi terkait. Kerangka tersebut, seperti yang digambarkan pada Gambar 4, mendefinisikan enam langkah dan empat set standar untuk melakukan evaluasi yang baik terhadap program kesehatan masyarakat.

Logika yang merupakan kerangka evaluasi adalah bahwa evaluasi yang baik tidak hanya mengumpulkan bukti yang akurat dan menarik kesimpulan yang valid, tetapi menghasilkan hasil yang digunakan untuk membuat perbedaan. Untuk memaksimalkan peluang hasil evaluasi yang akan digunakan, Anda perlu membuat "pasar" sebelum Anda membuat "produk"—evaluasi. Anda menentukan pasar dengan fokus pada evaluasi pada pertanyaan yang paling menonjol, relevan, dan penting. Anda memastikan fokus evaluasi terbaik dengan memahami mana pertanyaan-pertanyaan tersebut sesuai dengan lanskap penuh dari deskripsi program Anda, dan terutama dengan memastikan bahwa Anda telah menentukan dan melibatkan pemangku kepentingan yang peduli dengan pertanyaan-pertanyaan ini dan ingin mengambil keputusan tindakan terhadap hasilnya.



Gambar 4. Kerangka Evaluasi (CDC, 1999)

Langkah-langkah dalam kerangka CDC diinformasikan oleh serangkaian standar untuk evaluasi. Standar ini bukan merupakan cara untuk melakukan evaluasi; tapi, mereka bekerja untuk memandu pilihan Anda di antara banyak opsi yang tersedia di setiap langkah dalam kerangka. 30 standar menjadi empat kelompok:

Utilitas : Siapa yang membutuhkan hasil evaluasi? evaluasi memberikan informasi yang relevan secara tepat waktu bagi mereka?

Kelayakan : Apakah evaluasi waktu yang direncanakan realistis dengan mempertimbangkan, sumber daya, dan keahlian yang ada?

Kepentingan : Apakah evaluasi hak-hak individu dan kesejahteraan mereka yang terlibat? Apakah melibatkan mereka yang paling terkena dampak langsung program dan perubahan program, seperti peserta atau masyarakat sekitar?

Akurasi: evaluasi evaluasi temuan yang valid dan andal, mengingat kebutuhan mereka yang akan menggunakan hasilnya?

6.2. Bagaimana Membentuk Tim Evaluasi dan Memilih Lead Evaluator

Evaluasi yang baik membutuhkan kombinasi keterampilan yang jarang ditemukan pada satu orang. Pendekatan yang lebih disukai adalah memilih tim evaluasi yang mencakup staf program internal, pemangku kepentingan eksternal, dan mungkin konsultan atau kontraktor dengan evaluasi evaluasi.

Langkah awal dalam pembentukan adalah memutuskan siapa yang akan bertanggung jawab untuk merencanakan dan melaksanakan evaluasi. Program satu orang staf harus dipilih sebagai evaluator utama untuk mengoordinasikan program upaya. Orang ini harus bertanggung jawab untuk kegiatan evaluasi, termasuk perencanaan dan penganggaran untuk evaluasi, mengembangkan tujuan program, menangani kebutuhan pengumpulan data, melaporkan temuan, dan bekerja dengan konsultan. Memimpin evaluator pada akhirnya bertanggung jawab untuk melibatkan pemangku kepentingan, konsultan, dan kolaborator lain yang membawa keterampilan dan minat yang dibutuhkan untuk merencanakan dan melakukan evaluasi.

Meskipun orang staf ini harus memiliki keterampilan yang diperlukan untuk mengoordinasikan evaluasi secara kompeten, dia dapat memilih untuk mencari keahlian teknis di tempat lain untuk merancang dan mengimplementasikan tugas-tugas tertentu. Namun, pengembangan keahlian dan evaluasi evaluasi internal adalah tujuan yang bermanfaat bagi sebagian besar organisasi kesehatan masyarakat. Dari karakteristik evaluator yang baik yang tercantum dalam kotak teks di bawah ini, kemampuan evaluator untuk bekerja dengan kelompok pemangku kepentingan yang beragam perlu disoroti. Pemimpin evaluator harus mau dan mampu menarik keluar dan mendamaikan perbedaan nilai dan standar di antara pemangku kepentingan dan untuk bekerja dengan perwakilan pemangku kepentingan yang berpengetahuan luas dalam merancang dan melakukan evaluasi.

Mencari keahlian evaluasi tambahan dalam program dalam departemen kesehatan, melalui mitra eksternal (misalnya, universitas, organisasi, perusahaan), dari rekan program di negara bagian dan lokalitas lain, dan melalui bantuan teknis yang ditawarkan oleh pemerintah.

Anda juga dapat menggunakan konsultan luar sebagai sukarelawan, anggota panel, atau kontraktor. Konsultan Eksternal dapat memberikan penilaian evaluasi tingkat tinggi dari sudut pandang objektif. Faktor penting yang perlu dipertimbangkan ketika memilih konsultan adalah tingkat pelatihan profesional, pengalaman, dan kemampuan mereka untuk memenuhi kebutuhan Anda. Secara total, penting untuk

menemukan konsultan pendekatan evaluasi, latar belakang, dan pelatihannya paling sesuai dengan kebutuhan dan tujuan evaluasi program Anda. Pastikan untuk memeriksa semua referensi dengan cermat sebelum Anda membuat kontrak dengan konsultan mana pun.

Untuk menghasilkan diskusi seputar perencanaan dan evaluasi, beberapa negara bagian telah membentuk panel evaluasi. Panel, biasanya menghasilkan masukan dari lokal, regional atau pakar nasional sebaliknya sulit diakses. Panel perhatian Anda akan memberikan hasil pada upaya Anda dan terbukti berguna dalam dukungan luas untuk evaluasi.

Anggota tim evaluasi harus dengan jelas mendefinisikan peran mereka masing-masing. Konsensus informal mungkin cukup; yang lain lebih kesepakatan tertulis yang menjelaskan siapa yang akan melakukan evaluasi dan memberikan peran dan tanggung jawab khusus kepada anggota tim individu. berita, tim harus mengklarifikasi dan mencapai konvensi tentang:

- Tujuan Evaluasi
- Potensi pengguna evaluasi evaluasi dan rencana diseminasi Pendekatan Evaluasi
- Sumber daya yang tersedia
- Perlindungan bagi subyek manusia.

Bab 7.

LANGKAH-LANGKAH EVALUASI PROGRAM KESEHATAN MASYARAKAT

7.1. Langkah 1: Libatkan Pemangku Kepentingan

Pemangku kepentingan adalah orang atau organisasi yang berinvestasi dalam program, tertarik pada hasil evaluasi, dan/atau kepentingan dengan apa yang akan dilakukan dengan hasil evaluasi.

Mewakili kebutuhan dan minat mereka selama proses adalah dasar untuk evaluasi program yang baik.

Pemangku kepentingan utama untuk evaluasi program kesehatan masyarakat dibagi menjadi tiga kelompok besar:

1. Mereka yang terlibat dalam operasi program: Manajemen, staf program, mitra, dana lembaga, dan anggota.
2. Mereka yang dilayani atau terpengaruh oleh program: Pasien atau klien, kelompok advokasi, komunitas anggota, dan pejabat terpilih.
3. Mereka yang bermanfaat sebagai pengguna temuan temuan: Orang-orang dalam posisi untuk membuat keputusan tentang program, seperti mitra, sumber dana, masyarakat, dan umum atau pembayar pajak.

Jelas, kategori-kategori ini tidak saling eksklusif; Khususnya, pengguna utama temuan evaluasi sering kali adalah anggota dari dua kelompok lainnya, yaitu program manajemen atau organisasi atau advokasi. Meskipun Anda mungkin berpikir Anda mengenal pemangku kepentingan Anda dengan baik, kategori berikut membantu Anda untuk berpikir secara luas dan inklusif dalam mengidentifikasi pemangku kepentingan.

Pemangku kepentingan dapat membantu (atau menahan) evaluasi sebelum dilakukan, saat dilakukan, dan setelah hasilnya dikumpulkan dan siap digunakan. karena begitu banyak upaya kesehatan masyarakat yang kompleks dan karena lembaga kesehatan masyarakat beberapa penghapusan dari pelaksanaan garis depan, pemangku kepentingan mengambil kepentingan khusus dalam memastikan bahwa pertanyaan evaluasi yang tepat diidentifikasi dan hasil evaluasi akan digunakan untuk membuat perbedaan.

Pemangku kepentingan jauh lebih mungkin untuk mendukung evaluasi dan

bertindak berdasarkan hasil dan rekomendasi jika mereka terlibat dalam proses evaluasi. Sebaliknya, tanpa dukungan pemangku kepentingan, evaluasi Anda dapat diabaikan, dikritik, ditentang, atau bahkan disabotase.

Gunakan standar evaluasi untuk membantu mengidentifikasi pemangku kepentingan yang paling penting. Berikan prioritas kepada para pemangku kepentingan yang:

- Dapat meningkatkan upaya atau evaluasi Anda
- Jawaban jawab atas implementasi sehari-hari kegiatan yang merupakan bagian dari program
- Akan mengadvokasi atau mengubah program yang mungkin direkomendasikan oleh evaluasi
- Akan mengizinkan atau mengizinkan kelanjutan atau program perluasan.

Selain itu, agar tepat/ etis dan akurat, Anda perlu menyertakan mereka yang berpartisipasi dalam program dan terpengaruh oleh program atau evaluasinya.

Tabel 5. Contoh Pemangku Kepentingan

Siapa pemangku kepentingan utama yang kita pesan untuk:			
Tingkatkan upaya kami	Intervensi yang sentral untuk usaha ini	Advokasi perubahan untuk melembagakan usaha ini	Kelanjutan dana/ otorisasi atau ekspansi dari upaya ini
Asosiasi dokter Asosiasi komunitas	departemen kesehatan negara bagian dan lokal Otoritas perumahan	Kelompok advokasi Kelompok kesehatan ibu dan anak Asosiasi dokter masyarakat asosiasi	Legislator dan pembuat kebijakan di Tingkat federal dan negara bagian CDC Industri swasta Sistem pengadilan

Selama proses perencanaan evaluasi, Anda akan menanyakan beberapa atau semua pemangku kepentingan pertanyaan berikut:

Siapa yang Anda wakili dan mengapa Anda tertarik dengan program ini?

- Apa pentingnya program ini bagi Anda?
- Apa yang ingin Anda capai dari program ini?
- Kejutan besar kemajuan yang Anda harapkan dari program ini saat ini?
- Apa yang Anda lihat sebagai pertanyaan evaluasi kritis saat ini?
- Bagaimana Anda akan menggunakan hasil evaluasi ini?
- Sumber daya apa (yaitu, waktu, dana, evaluasi, akses ke responden, dan akses kepada pembuat kebijakan)
- Dapatkah Anda berkontribusi pada upaya evaluasi ini?

Perspektif pemangku kepentingan dapat mempengaruhi setiap langkah CDC. Para pemangku kepentingan dalam “menggambarkan program” memastikan pemahaman yang jelas dan konsensual tentang kegiatan dan hasil program. Ini adalah latar belakang penting untuk memasukkan pemangku kepentingan yang lebih berharga dalam “memfokuskan desain evaluasi” untuk memastikan bahwa pertanyaan kunci yang paling penting akan dimasukkan.

Pemangku kepentingan juga mungkin memiliki wawasan atau preferensi tentang cara yang paling efektif dan tepat untuk mengumpulkan data dari responden sasaran. Dalam “membenarkan kesimpulan”, perspektif dan nilai yang dibawa ke proyek secara eksplisit diakui dan pemangku kepentingan dalam membuat bukti tentang yang dikumpulkan. Akhirnya, banyak waktu dan upaya yang dihabiskan dalam melibatkan dan membangun konvensi di antara para pemangku kepentingan terbayar pada langkah-langkah terakhir, karena keterlibatan pemangku kepentingan menciptakan pasar untuk hasil evaluasi.

Pemangku kepentingan dapat dilibatkan dalam evaluasi di berbagai tingkatan. Misalnya, Anda mungkin ingin memasukkan anggota ke dalam tim evaluasi dan melibatkan mereka dalam pengembangan, pengumpulan data, dan analisis. Atau pertimbangkan cara untuk menilai kebutuhan dan mitra Anda dalam evaluasi, dan kembangkan cara agar mereka tetap mendapat tentang kemajuannya dan mengintegrasikan ide-ide mereka ke dalam kegiatan evaluasi. Sekali lagi, pemangku kepentingan lebih mungkin mendukung evaluasi dan bertindak berdasarkan hasil dan rekomendasi jika mereka terlibat dalam proses evaluasi.

Libatkan juga kritikus program Anda dalam evaluasi. Kritik dapat membantu mengidentifikasi masalah seputar program strategi Anda dan evaluasi

informasi yang dapat diserang atau didiskreditkan, sehingga membantu Anda memperkuat proses evaluasi. Informasi ini mungkin juga membantu Anda dan orang lain memahami alasan oposisi dan dapat membantu Anda melibatkan agen perubahan potensial di oposisi. Namun, berhati-hatilah: Penting untuk memahami motif oposisi sebelum melibatkan mereka dengan cara apa pun yang berarti.

Penekanan pada pelibatan pemangku kepentingan ini mencerminkan atau menonjolkannya model partisipatif penelitian “aksi” dalam komunitas riset. Pendekatan partisipatif yang menggabungkan penyelidikan dengan para pemangku kepentingan yang beragam untuk memenuhi kebutuhan khusus dan menghadapi masalah dan keadilan yang luas.

Tabel 6. Standar untuk Langkah 1

Standard	Questions
Utility	▪ Who will use these results?
Feasibility	▪ How much time and effort can be devoted to stakeholder engagement?
Propriety	▪ Which stakeholders need to be consulted to conduct an ethical evaluation, for example, to ensure we will identify negative as well as positive aspects of the program?
Accuracy	▪ How broadly do we need to engage stakeholders to paint an accurate picture of this program?

Tabel 7. Langkah untuk Melibatkan Pemangku Kepentingan

Checklist for Step 1: Engage Stakeholders	
<input type="checkbox"/>	Identify stakeholders, using the three broad categories discussed: those affected, those involved in operations, and those who will use the evaluation results.
<input type="checkbox"/>	Review the initial list of stakeholders to identify key stakeholders needed to improve credibility, implementation, advocacy, or funding/authorization decisions.
<input type="checkbox"/>	Engage individual stakeholders and/or representatives of stakeholder organizations
<input type="checkbox"/>	Create a plan for stakeholder involvement and identify areas for stakeholder input..
<input type="checkbox"/>	Target selected stakeholders for regular participation in key steps, including writing the program description, suggesting evaluation questions, choosing evaluation questions, and disseminating evaluation result

7.2. Langkah 2: Penjelasan Program

Sebuah deskripsi program yang komprehensif menjelaskan semua komponen dan hasil yang diinginkan dari program, sehingga membantu Anda berfokus pada evaluasi Anda pada pertanyaan yang paling sentral dan penting. Perhatikan bahwa dalam langkah ini Anda menggambarkan program dan bukan evaluasi. Dalam bab ini, Anda akan menggunakan alat yang disebut "pemodelan logika" untuk menggambarkan komponen program ini, tetapi deskripsi program dapat dikembangkan tanpa menggunakan alat ini atau alat apa pun.

Langkah ini dapat mengikuti langkah pemangku kepentingan atau mempersiapkannya. Bagaimanapun, kombinasi keterlibatan pemangku kepentingan dan deskripsi program mempromosikan dan konvensi sebelum data tersedia untuk mengukur efektivitas program. Kejelasan ini menetapkan tahap untuk evaluasi program yang baik, dan, juga dapat membantu memastikan bahwa perencanaan strategi dan pengukuran kinerja, beroperasi dari kerangka yang sama tentang program tersebut.

Deskripsi program yang komprehensif mencakup komponen-komponen berikut:

- **Kebutuhan** . Apa masalah kesehatan masyarakat besar yang ingin Anda atasi dengan program Anda?

- **Sasaran** . Kelompok atau organisasi mana yang perlu diubah atau mengambil tindakan untuk memastikan kemajuan masalah kesehatan masyarakat?
- **Hasil** . Bagaimana dan dengan cara apa target-target ini perlu diubah? Tindakan khusus yang mereka pesan untuk diambil?
- **Kegiatan** . Apa yang akan dilakukan program Anda dan stafnya untuk menggerakkan kelompok sasaran ini untuk berubah/mengambil tindakan?
- **Keluaran** . Kapasitas atau produk nyata apa yang akan dihasilkan oleh program Anda? kegiatan?
- **Sumber Daya/Input** . Apa yang dibutuhkan dari lingkungan yang lebih besar agar kegiatan berhasil dipasang?
- **Hubungan Kegiatan dan Hasil** . Kegiatan mana yang sedang dilaksanakan untuk menghasilkan kemajuan pada hasil yang mana?

Selain menentukan komponen-komponen ini, deskripsi program yang mencakup diskusi lengkap tentang:

- **Tahap Perkembangan** . Apakah programnya baru saja dimulai, apakah dalam implementasinya? tahap, atau telah berlangsung untuk jangka waktu yang signifikan?
- **Konteks** . Faktor dan tren apa di lingkungan yang lebih besar yang dapat memengaruhi program? sukses atau gagal?

Setelah komponen dari deskripsi program telah diidentifikasi, penggambaran visual dapat membantu untuk meringkas hubungan antar komponen. Kejelasan ini dapat membantu perencanaan strategi dan evaluasi program. Meskipun ada cara lain untuk menggambarkan hubungan ini, model logika adalah alat yang umum digunakan oleh evaluator, dan alat tersebut dijelaskan paling lengkap dalam kerangka CDC.

Model logika adalah penggambaran grafis dari hubungan antara kegiatan dan hasil yang diharapkan. Dua kata dalam definisi ini:

Hubungan : Model logika menambahkan tidak hanya kegiatan yang membentuk program dan hubungan antar kegiatan tersebut, tetapi hubungan antara komponen dan hasil.

Hasil yang diharapkan : Model menggambarkan hasil yang diinginkan dari program kegiatan. Sebagai awal untuk evaluasi dan perencanaan, model tersebut berfungsi sebagai “peta jalan hasil” yang menunjukkan

logika titik balik program, yaitu mengapa program itu harus berhasil. Artinya, dari semua kegiatan yang bisa dilakukan untuk mengatasi masalah ini, kegiatan ini dipilih karena, jika dilaksanakan sebagaimana direncanakan akan mengarah pada hasil yang diinginkan. Seiring waktu, evaluasi, penelitian, dan pengalaman sehari-hari akan memperdalam pemahaman tentang apa yang berhasil dan tidak, dan modelnya akan berubah sesuai dengan itu.

Model logika tidak memerlukan pemikiran baru tentang program; mengubah bahan mentah yang dihasilkan dalam deskripsi program menjadi gambar program.

Tabel 8. Standar untuk Langkah 2: Penjelasan Program

Standar	Pertanyaan
Kegunaan	<p>Memikirkan bagaimana model akan digunakan, apakah tingkat detailnya sesuai atau ada terlalu banyak atau terlalu sedikit detail?</p> <p>Apakah keputusan /deskripsi program dapat dijangkau oleh mereka yang perlu digunakan untuk membuat perencanaan evaluasi?</p>
Kelayakan	<p>Apakah deskripsi program mencakup setidaknya beberapa kegiatan dan hasil yang mengendalikan program?</p>
Kepatutan	<p>Apakah evaluasi lengkap dan adil dalam menilai semua aspek program, termasuk kekuatan dan kelemahannya?</p> <p>Apakah deskripsi program mencakup cukup detail untuk memeriksa kekuatan kedua? dan kelemahan, dan hasil yang tidak diinginkan serta yang diinginkan?</p>
Ketepatan	<p>Apakah deskripsi program komprehensif?</p> <p>Apakah Anda mengutamakan mendokumentasikan konteks program sehingga kemungkinan memengaruhi program dapat diidentifikasi?</p>

Tabel 9. Daftar Periksa untuk Langkah 2

Checklist for Step 2: Describing the Program	
<input type="checkbox"/>	Compile a comprehensive program description including need, targets, outcomes, activities, and resources.
<input type="checkbox"/>	Identify the stage of development and context of the program.
<input type="checkbox"/>	Convert inputs, activities, outputs, and outcomes into a simple global logic model.
<input type="checkbox"/>	Elaborate the model as needed.
<input type="checkbox"/>	Develop more detailed models from the global model as needed.

7.3 Langkah 3: Disain Evaluasi

Setelah menyelesaikan Langkah 1 dan 2, Anda dan pemangku kepentingan Anda harus memiliki pemahaman yang jelas tentang program dan mencapai konvensi. Sekarang tim evaluasi Anda perlu difokuskan evaluasi. Ini termasuk menentukan pertanyaan evaluasi yang paling penting dan desain yang tepat untuk evaluasi.

Memfokuskan evaluasi mengasumsikan bahwa seluruh program tidak perlu dimulai setiap saat. Sebaliknya, evaluasi program yang tepat tergantung pada pertanyaan apa yang diajukan, siapa yang mengajukan pertanyaan, dan apa yang akan dilakukan dengan informasi tersebut.

Karena sumber daya untuk evaluasi selalu terbatas, bab ini memberikan penilaian keputusan untuk membantu Anda menentukan fokus evaluasi terbaik kapan saja. Kriteria ini diilhami oleh standar evaluasi: khususnya, utilitas (siapa yang akan menggunakan hasil dan informasi yang paling berguna bagi mereka) dan ketersediaan (berapa banyak waktu dan sumber daya yang tersedia untuk evaluasi).

Model logika yang dikembangkan pada Langkah 2 mengatur tahapan untuk menentukan fokus evaluasi terbaik. Pendekatan evaluasi dalam Kerangka Evaluasi CDC sedikit dari pendekatan evaluasi tradisional.

Daripada evaluasi sumatif, dilakukan ketika program berjalan dan bertanya "Apakah program itu berhasil?" kerangka kerja CDC memandang evaluasi sebagai aktivitas berkelanjutan selama program yang menanyakan, "Apakah program itu berhasil?"

Oleh karena itu, sebuah program selalu siap untuk beberapa evaluasi. Karena model menampilkan program dari input melalui aktivitas/output hingga urutan dari jangka pendek hingga paling jauh, model ini dapat memandu diskusi tentang apa yang dapat Anda harapkan untuk dicapai pada titik tertentu dalam proyek kehidupan Anda. Anda fokus pada hasil distal, atau hanya pada jangka pendek atau menengah? Atau sebaliknya, apakah evaluasi proses paling masuk akal saat ini?

JENIS EVALUASI

Banyak pertanyaan yang berbeda dapat menjadi bagian dari evaluasi program; tergantung pada berapa lama program telah ada, siapa yang mengajukan pertanyaan, dan mengapa evaluasi diperlukan. Secara umum, pertanyaan

evaluasi untuk program yang ada¹⁷ termasuk dalam salah satu kelompok berikut:

1. Evaluasi Proses

Evaluasi pelaksanaan (evaluasi proses) mendokumentasikan apakah suatu program telah dilaksanakan sebagaimana dimaksud—dan mengapa mengapa tidak? Dalam evaluasi, Anda dapat memeriksa apakah kegiatan sedang berlangsung, siapa yang melakukan kegiatan, siapa saja yang dicapai melalui kegiatan, dan apakah masukan yang telah dialokasikan atau dimobilisasi.

Proses evaluasi penting untuk membantu membedakan penyebab kinerja program yang buruk—apakah program itu ide yang buruk, atau apakah ide yang bagus tidak dapat mencapai standar implementasi yang Anda tetapkan? Dalam semua kasus, evaluasi proses mengukur apakah kinerja program aktual sesuai dengan rencana awal. Pengukuran tersebut mungkin termasuk kontras kinerja aktual dan yang direncanakan sepanjang semua atau beberapa hal berikut:

- Lokasi di mana layanan atau program yang disediakan (misalnya, pedesaan, perkotaan)
- Jumlah orang yang menerima layanan
- Status ekonomi dan latar belakang ras/etnis orang yang menerima layanan
- Kualitas pelayanan
- Peristiwa aktual yang terjadi saat jasa diberikan
- Jumlah uang yang digunakan proyek
- Pendanaan langsung dan dalam bentuk barang untuk layanan
- Staf untuk layanan atau program
- Jumlah kegiatan dan pertemuan
- Jumlah sesi pelatihan yang dilakukan

Ketika sumber daya evaluasi terbatas, hanya masalah implementasi yang paling penting yang dapat dimasukkan.

Berikut adalah beberapa "hal lazim" yg mengkompromikan implementasi dan mungkin dipertimbangkan untuk dimasukkan dalam evaluasi proses:

- Transfer Akuntabilitas: Ketika program kegiatan tidak dapat menghasilkan hasil yang diinginkan kecuali beberapa orang atau organisasi lain mengambil tindakan yang tepat, ada akuntabilitas transfer.

- Ukuran : Hasil yang diharapkan dari program kegiatan (misalnya, pelatihan, manajemen kasus, konseling) dapat mengasumsikan tingkat partisipasi atau paparan campur tangan.
- Akses : Ketika hasil yang diinginkan tidak hanya membutuhkan peningkatan permintaan konsumen tetapi juga peningkatan pasokan layanan untuk memenuhinya, maka evaluasi dapat dicapai dengan langkah-langkah akses.
- Kompetensi Staf: Hasil yang diinginkan mungkin dirancang dengan baik program kegiatan yang disampaikan oleh staf yang tidak hanya kompeten secara teknis tetapi juga sesuai dengan audiens target. Ukuran kecocokan staf dan target audiens mungkin dimasukkan dalam proses evaluasi.

2. Evaluasi Hasil/ Dampak

Evaluasi kemajuan pada urutan hasil yang akan ditangani oleh program. Program sering menggambarkan urutan ini menggunakan istilah seperti hasil jangka pendek, menengah, dan jangka panjang, atau proksimal (dekat dengan intervensi) atau distal (jauh dari intervensi).

Tergantung pada tahap pengembangan program dan tujuan evaluasi, evaluasi hasil dapat mencakup satu atau semua hasil dalam urutan, termasuk:

- Perubahan sikap dan keyakinan masyarakat
- Perubahan risiko atau perilaku protektif
- Perubahan lingkungan, termasuk kebijakan publik dan swasta, penegakan peraturan formal dan informal, dan pengaruh norma sosial dan kekuatan sosial lainnya
- Perubahan tren morbiditas dan mortalitas

Sementara proses dan hasil adalah yang paling umum, ada beberapa evaluasi yang penting untuk evaluasi program tertentu. Ini termasuk yang berikut:

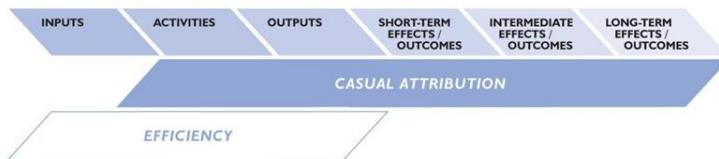
- Efisiensi : Apakah program yang Anda hasilkan dengan penggunaan sumber daya yang minimal seperti anggaran dan waktu staf? Berapa volume keluaran yang dihasilkan oleh sumber daya yang dicurahkan untuk program Anda?
- Efektivitas Biaya: Apakah nilai atau manfaat dari hasil program Anda melebihi biaya produksinya?

- Atribusi : Bisa hasil terkait dengan program Anda, bukan dengan hal lain yang terjadi pada waktu yang sama?

Semua jenis pertanyaan evaluasi ini berhubungan dengan sebagian, tetapi tidak semua, dari model logika. Gambar 5 dan Gambar 6 menunjukkan di mana dalam model logika setiap jenis evaluasi akan fokus.



Gambar 5.



Gambar 6.

Implementasi evaluasi akan berfokus pada masukan, kegiatan, dan keluaran dan tidak memperhatikan kinerja pada hasil. Evaluasi keefektifan akan melakukan sebaliknya—berfokus pada beberapa atau semua kotak hasil, tetapi tidak harus pada aktivitas yang menghasilkannya.

Evaluasi efisiensi memperhatikan panah yang menghubungkan input dengan aktivitas/output—berapa banyak output yang dihasilkan untuk level input/sumber daya tertentu.

Atribusi akan fokus pada panah antara/ keluaran spesifik dan hasil spesifik—apakah kemajuan pada hasil terkait dengan aktivitas/keluaran spesifik.

Menentukan fokus evaluasi yang benar adalah keputusan kasus per kasus. Beberapa pedoman yang diilhami oleh standar evaluasi “utilitas” dan “kelayakan” dapat membantu menentukan fokus terbaik.

Pertimbangan UTILITAS

1. Apa tujuan evaluasi?

Tujuan mengacu pada tujuan umum dari evaluasi. Tujuan yang jelas berfungsi sebagai dasar untuk pertanyaan evaluasi, desain, dan metode.

Beberapa tujuan umum:

- Memperoleh pengetahuan baru tentang program
- Peningkatan atau penyempurnaan program yang ada (misalnya, proses atau program strategi)
- Menentukan efek dari suatu program dengan memberikan bukti mengenai kontribusi program untuk tujuan jangka panjang
- Mempengaruhi peserta program dengan bertindak sebagai katalis untuk perubahan yang diarahkan sendiri (misalnya, mengajar)

2. Siapa yang akan menggunakan hasil evaluasi?

Pengguna adalah individu atau organisasi yang akan menggunakan temuan evaluasi. Pengguna kemungkinan besar telah diidentifikasi selama Langkah 1 dalam proses pelibatan pemangku kepentingan. Pada langkah-langkah ini, Anda perlu memasukkan masukan mereka ke dalam desain evaluasi dan pertanyaan evaluasi. Dukungan dari pengguna yang dituju akan meningkatkan kemungkinan bahwa hasil evaluasi akan digunakan untuk perbaikan program.

3. Bagaimana mereka akan menggunakan hasil evaluasi?

Banyak wawasan tentang penggunaan akan diidentifikasi pada Langkah 1. Informasi yang dikumpulkan mungkin memiliki kegunaan yang berbeda-beda, yang harus dijelaskan secara rinci saat merencanakan evaluasi. Beberapa contoh penggunaan informasi evaluasi:

- Untuk mendokumentasikan tingkat keberhasilan dalam mencapai tujuan
- Untuk mengidentifikasi program daerah yang perlu ditingkatkan
- Untuk memutuskan alokasi sumber daya
- Memobilisasi dukungan masyarakat
- Untuk mendistribusikan kembali atau perluasan lokasi di mana intervensi dilakukan
- Untuk meningkatkan isi program
- Untuk fokus pada program sumber daya pada populasi tertentu Untuk meminta lebih banyak dana atau mitra tambahan

4) Apa yang dibutuhkan pemangku kepentingan utama lainnya dari evaluasi ?

Tentu saja, pemangku kepentingan yang paling penting adalah mereka yang meminta atau yang akan menggunakan hasil evaluasi. Namun demikian,

pada Langkah 1, Anda mungkin juga telah mengidentifikasi pemangku kepentingan yang, meskipun tidak menggunakan evaluasi saat ini, memiliki pertanyaan kunci yang mungkin perlu ditangani dalam evaluasi agar mereka tetap terlibat. Misalnya, pemangku kepentingan tertentu mungkin selalu khawatir tentang biaya, disparitas, atau atribusi. Jika demikian, Anda mungkin perlu menambahkan pertanyaan-pertanyaan itu ke fokus evaluasi Anda.

PERTIMBANGAN KELAYAKAN

Empat pertanyaan pertama membantu mengidentifikasi fokus evaluasi yang paling berguna, tetapi Anda juga harus menentukan apakah itu realistis/layak. Tiga pertanyaan memberikan pemeriksaan realitas pada fokus yang Anda inginkan:

1) Bagaimana tahapan pengembangan program?

Selama Langkah 2, Anda akan menentukan tahap pengembangan program. Ada kira-kira tiga tahap dalam pengembangan program --perencanaan, implementasi, dan pemeliharaan- yang menunjukkan fokus yang berbeda.

Pada tahap perencanaan, evaluasi yang benar-benar formatif—siapa target Anda, bagaimana Anda menjangkau mereka, berapa biayanya—mungkin menjadi fokus yang paling tepat sebuah evaluasi yang mencakup hasil tidak akan masuk akal pada tahap ini. Sebaliknya, evaluasi program dalam tahap pemeliharaan perlu memasukkan beberapa pengukuran kemajuan hasil, bahkan jika itu juga mencakup implementasi pengukuran.

2) Mungkinkah mengintensifkan program?

Beberapa program yang sangat luas dan beragam. Orang lain mungkin hanya menggunakan satu pendekatan

untuk mengatasi masalah besar. Beberapa program memberikan paparan ekstensif (“dosis”) program, sementara yang lain melibatkan peserta dengan cepat dan dangkal. Program sederhana atau dangkal, meskipun mungkin bermanfaat, tidak dapat realistis diharapkan memberikan kontribusi yang signifikan terhadap hasil akhir dari program yang lebih besar, bahkan ketika program tersebut beroperasi penuh.

3) Apa pertimbangan sumber daya dan logistik yang relevan

Sumber daya dan logistik dapat mempengaruhi keputusan tentang fokus evaluasi. Beberapa hasil lebih cepat, lebih mudah, dan lebih murah untuk diukur,

sementara yang lain mungkin tidak dapat diukur sama sekali. Fakta-fakta ini dapat memiringkan keputusan tentang fokus terhadap evaluasi beberapa hasil yang bertentangan dengan yang lain.

Identifikasi awal ketidakkonsistenan antara utilitas dan prioritas merupakan bagian penting dari langkah evaluasi. Tetapi kita juga harus memastikan “pertemuan pikiran” tentang apa yang merupakan fokus realistis untuk evaluasi program setiap saat.

Contoh perumahan yang terjangkau menunjukkan bagaimana fokus yang diinginkan mungkin dibatasi oleh kenyataan. Model logika yang diuraikan penting dalam kasus ini. Ini menjelaskan bahwa, sementara staf program fokus pada produksi rumah baru, pemangku kepentingan penting seperti organisasi berbasis masyarakat dan donor berbasis agama berkomitmen untuk hasil yang lebih jauh seperti perubahan dalam hasil kehidupan keluarga, atau pada hasil investasi luar dalam Komunitas. Model tersebut mengarah pada diskusi tentang ekspektasi yang masuk akal dan, pada akhirnya, untuk evaluasi evaluasi yang mencakup beberapa hasil yang lebih jauh, yang mengarah pada apresiasi pemangku kepentingan yang lebih besar dari menuju hasil yang mereka inginkan.

MEMUTUSKAN DESAIN EVALUASI

Untuk menentukan fokus evaluasi dan pertanyaan evaluasi yang spesifik, pada titik ini Anda juga perlu menentukan desain evaluasi yang sesuai. Kepentingan utama dalam memilih desain evaluasi adalah apakah Anda diminta untuk menyelesaikan kemajuan atau apakah Anda juga diminta untuk menunjukkan atribusi— bahwa kemajuan hasil terkait dengan upaya program Anda. Pertanyaan atribusi mungkin lebih tepat dilihat sebagai penelitian sebagai lawan evaluasi program, tergantung pada tingkat pengawasan yang dibuat.

Tiga jenis desain umum penelitian yang umum dikenal: eksperimental, kuasi eksperimental, dan non eksperimental/ observasional. Evaluasi program tradisional biasanya menggunakan tipe ketiga, tetapi ketiganya disajikan di sini, selama program berlangsung, pendekatan evaluasi tradisional mungkin perlu dilengkapi dengan penelitian lain yang lebih mirip dengan penelitian.

Desain eksperimental menggunakan penugasan acak untuk membandingkan hasil intervensi pada satu atau lebih kelompok dengan kelompok setara atau kelompok yang tidak menerima.

Misalnya, Anda dapat memilih sekelompok sekolah yang serupa, lalu secara acak menetapkan beberapa sekolah untuk menerima kurikulum pencegahan dan

sekolah lain sebagai kontrol. Semua sekolah memiliki kesempatan yang sama untuk dipilih sebagai sekolah intervensi atau sekolah kontrol. Penugasan acak, mengurangi kemungkinan bahwa sekolah kontrol dan intervensi bervariasi dengan cara apa pun yang dapat mempengaruhi perbedaan dalam hasil program. Ini mendukung Anda untuk menghubungkan perubahan hasil dengan program Anda. Misalnya, jika siswa di sekolah intervensi menunda onset atau perilaku berisiko lebih lama daripada siswa di sekolah kontrol, Anda dapat menghubungkan keberhasilan program Anda. Namun, dalam pengaturan komunitas, sulit, atau bahkan terkadang tidak etis, untuk memiliki kelompok kontrol yang sebenarnya.

Meskipun ada beberapa solusi yang menjaga integritas desain eksperimental, pilihan lain adalah menggunakan desain kuasi-eksperimental. Desain ini membuat perbandingan antara kelompok nonequivalent dan tidak melibatkan penugasan acak untuk intervensi dan kelompok kontrol.

Contohnya adalah keyakinan orang dewasa tentang akibat berbahaya dari asap tembakau lingkungan (ETS) di dua komunitas, kemudian melakukan kampanye media di salah satu komunitas. Setelah kampanye, Anda akan menilai kembali orang dewasa dan berharap menemukan persentase orang dewasa yang lebih tinggi yang percaya bahwa ETS berbahaya di komunitas yang menerima kampanye media. Kritikus dapat berargumen bahwa perbedaan lain antara komunitas kedua menyebabkan perubahan keyakinan, jadi penting untuk mendokumentasikan bahwa intervensi dan kelompok pembanding serupa pada faktor-faktor kunci seperti demografi dan arus atau sejarah terkait.

Desain kuasi-eksperimental, perbandingan data hasil/hasil antar negara dan antara satu negara bagian dengan bangsa secara keseluruhan adalah cara umum untuk eksperimen upaya kesehatan masyarakat. Perbandingan itu akan membantu Anda menetapkan tolok ukur yang berarti untuk kemajuan.

Negara bagian perbandingan dapat kemajuan mereka dengan bagian dengan investasi serupa di bidang kesehatan masyarakat, atau mereka dapat membandingkan hasil mereka dengan hasil yang diharapkan jika program serupa dengan negara bagian dengan investasi yang lebih besar.

Perbandingan data juga berguna untuk mengukur indikator dalam mengantisipasi program baru atau memperluas. Misalnya, mencatat perubahan dalam indikator kunci waktu sebelum implementasi program menunjukkan perlunya program Anda dan membantu kemajuan komparatif negara dengan program kesehatan masyarakat yang sudah ada. Kurangnya perubahan dalam indikator dapat berguna sebagai pembenaran untuk investasi yang lebih besar

dalam program berbasis bukti, pendanaan dengan baik, dan lebih komprehensif. Perbandingan antar keadaan dapat disorot dengan analisis deret waktu. Misalnya, pertanyaan tentang banyak sistem pengawasan nasional yang lebih besar tidak akan berubah dalam beberapa tahun, sehingga Anda dapat membuat perbandingan dengan negara bagian lain dari waktu ke waktu, dengan menggunakan indikator tertentu.

Berkolaborasi dengan ahli epidemiologi negara bagian, koordinator pengawasan, dan ahli statistik untuk menjadikan perbandingan negara bagian dan nasional sebagai komponen penting dari evaluasi Anda.

Desain observasional termasuk, tetapi tidak terbatas pada, analisis deret waktu, survei cross-sectional, dan studi kasus. Survei berkala lintas bagian (misalnya, YTS atau BRFS) dapat memberitahukan evaluasi Anda. Studi kasus mungkin sangat tepat untuk menilai perubahan kapasitas kesehatan masyarakat dalam kelompok populasi yang berbeda. Studi kasus dapat diterapkan ketika programnya unik, ketika program yang digunakan dalam pengaturan yang berbeda, ketika hasil unik sedang dinilai, atau ketika lingkungan sangat tidak terduga. Studi kasus juga dapat mendukung eksplorasi karakteristik masyarakat dan bagaimana hal ini dapat mempengaruhi pelaksanaan program, serta identifikasi penghambat dan fasilitator perubahan.

Masalah "atribusi kausal" ini, walaupun sering menjadi pertanyaan penelitian utama, mungkin perlu atau tidak perlu melengkapi evaluasi program tradisional. Bidang kesehatan masyarakat berada di bawah tekanan yang meningkat untuk menunjukkan bahwa program itu bermanfaat, efektif, dan efisien. Selama dua dekade terakhir, pengetahuan dan pemahaman tentang bagaimana program yang kompleks meningkat secara signifikan. Namun demikian, karena program sangat kompleks, desain penelitian tradisional yang dijelaskan di sini mungkin bukan pilihan yang baik. Sebagai Organisasi Kesehatan Dunia mencatat, "penggunaan uji coba kontrol acak untuk mengungkap kasus-kasus promosi kesehatan, dalam banyak, tidak tepat, menyesatkan, dan tidak perlu mahal."

Sesuaikan dan sesuaikan desain yang kurang tradisional (misalnya, desain sederhana sebelum-sesudah [pretest-posttest] atau posttest-only). Tergantung pada tujuan program Anda dan tujuan penggunaan evaluasi, desain ini lebih cocok untuk mengukur kemajuan pencapaian tujuan program. Bahkan ketika ada kebutuhan untuk membuktikan bahwa program bertanggung jawab atas kemajuan hasil, desain tradisional mungkin bukan satu-satunya atau alternatif terbaik. Bergantung pada seberapa ketat pembuktiannya, kedekatan

waktu antara pelaksanaan program dan hasil, atau penghapusan pendekatan mungkin cukup untuk mendukung pemangku kepentingan utama bahwa program memberikan kontribusi. Meskipun alternatif desain ini sering kali lebih murah dan membutuhkan lebih sedikit waktu, perlu diingat bahwa menghemat waktu dan tidak boleh menjadi kriteria utama dalam memilih desain evaluasi. Penting untuk memilih desain yang akan mengukur apa yang perlu Anda ukur dan yang akan memenuhi kebutuhan langsung dan jangka panjang Anda.

Alternatif lain untuk model eksperimental dan kuasi-eksperimental adalah evaluasi berbasis tujuan model, yang menggunakan program tujuan yang telah ditentukan sebelumnya dan teori yang dibuat sebagai evaluasi standar, sehingga program bertanggung jawab terhadap harapan sebelumnya.

Penekanan Kerangka CDC pada deskripsi program dan pembangunan model logika menetapkan tanggung jawab untuk evaluasi program berbasis tujuan yang kuat. Dalam kasus seperti itu, evaluasi evaluasi berfokus pada kegiatan; keluaran; dan hasil jangka pendek, menengah, dan panjang yang digariskan dalam model logika program untuk mengarahkan kegiatan pengukuran.

Desain yang Anda pilih untuk memengaruhi waktu pengumpulan data, cara Anda menganalisis data, dan jenis kesimpulan yang dapat Anda buat dari temuan Anda. Pendekatan kolaboratif untuk berfokus pada evaluasi cara praktis untuk memastikan dan memanfaatkan desain evaluasi Anda dengan lebih baik.

Tabel 10. Standar untuk Langkah 3

Standar	Pertanyaan
Kegunaan	<p>Apa tujuan evaluasi?</p> <p>Sapa yang akan menggunakan hasil evaluasi dan bagaimana mereka akan menggunakannya?</p> <p>Apakah kebutuhan khusus dari pemangku kepentingan lain yang harus ditangani?</p>
Kelayakan	<p>Bagaimana tahap pengembangan program?</p> <p>Kejutan programnya?</p> <p>fokus pada komponen dalam fokus yang diusulkan?</p>
Kepatutan	<p>fokus fokus dan desain cukup mendeteksi yang tidak diinginkan?</p> <p>Apakah fokus dan desain pengalaman mereka yang terkena dampak program?</p>
Ketepatan	<p>Apakah fokusnya cukup luas untuk mendeteksi keberhasilan atau kegagalan program?</p> <p>Apakah desainnya tepat untuk menjawab pertanyaan—seperti atribusi—yang dibuat oleh pemangku kepentingan?</p>

7.4 Langkah 4: Pengumpulan Data

Pengumpulan bukti/ data harus mencakup pertimbangan berikut:

- Indikator
- Sumber bukti/metode pengumpulan data
- Kualitas
- Kuantitas
- Logistik

Tabel 11. Contoh Indikator Komponen Program dalam Evaluasi

Komponen Program	Indikator
Penyedia pelatihan	Serangkaian 3 pelatihan akan dilakukan di 4 wilayah negara bagian
Presentasi LHD pendidik perawat Pendidik perawat	akan membuat presentasi ke 10 departemen kesehatan lokal (LHD) terbesar
Putaran rekan dokter	Dokter akan menyelenggarakan putaran pendidikan sebaya di 10 rumah sakit terbesar
Penyedia menghadiri pelatihan dan pelatihan	Pelatihan akan dihadiri dengan baik dan mencerminkan perpaduan yang baik antara spesialisasi dan representasi geografis
Penyedia menerima dan menggunakan tool kit 50%	50%+ penyedia yang menerima tool kit akan melaporkan penggunaannya (atau kartu "ajakan bertindak" akan diterima dari 25% dari semua penyedia yang menerima tool kit)
Perawat LHD melakukan konsultasi penyedia swasta	Perawat di LHD akan melakukan konsultasi penyedia dengan penyedia terbesar di daerah
Penyedia KAB meningkat	Penyedia menunjukkan peningkatan pengetahuan, sikap, dan kredo (KAB) pada item utama yang dipilih
Motivasi penyedia meningkat	Niat penyedia untuk meningkat

Tabel 12. Contoh Metode Pengumpulan Data dan Sumber Indikator (Penyelenggara Program Edukasi Imunisasi)

Indikator	Metode/Sumber Pengumpulan Data
Serangkaian 3 pelatihan akan dilakukan di 4 wilayah negara bagian	Log pelatihan
Perawat pendidik akan membuat presentasi ke 10 departemen kesehatan lokal (LHD) terbesar	Log pelatihan
Dokter akan menyelenggarakan pelatihan pendidikan sebaya di 10 rumah sakit terbesar	Log pelatihan
Pelatihan akan diikuti dengan baik dan mencerminkan perpaduan yang baik antara spesialisasi dan representasi geografis	Informasi pendaftaran
50%+ penyedia yang menerima tool kit akan melaporkan penggunaannya (atau kartu "ajakan bertindak" akan diterima dari 25% dari semua penyedia yang menerima tool kit)	Penyedia pengawasan Analisis/jumlah kartu ajakan bertindak
Perawat di LHD akan melakukan konsultasi penyedia dengan praktik penyedia terbesar di kabupaten	Survei perawat, survei penyedia, atau log pelatihan
Penyedia menunjukkan peningkatan pengetahuan, sikap, dan kredo (KAB) pada item kunci yang dipilih	Survei penyedia, atau kelompok fokus, atau penyadapan
Niat penyedia untuk mengimunisasi meningkat	Survei penyedia, atau kelompok fokus, atau penyadapan

7.5. Langkah 5: Menarik Kesimpulan

Apakah Anda melakukan untuk menunjukkan efektivitas program, membantu meningkatkan program, atau menunjukkan akuntabilitas, Anda perlu

menganalisis dan memantau bukti yang dikumpulkan pada Langkah 4. Langkah 5 mencakup analisis bukti, membuat klaim berdasarkan analisis, dan pembuktian dengan membandingkan nilai-nilai pemangku kepentingan.

Mengapa langkah ini tidak disebut menganalisis data? Karena sama pentingnya dengan analisis data untuk evaluasi, evaluator tahu bahwa bukti yang dikumpulkan untuk evaluasi tidak selalu berbicara sendiri. Kesimpulan tidak didasarkan pada kesimpulan yang dianalisis dan disintesis (“bukti”) melalui prisma nilai (standar yang dibawa oleh pemangku kepentingan, dan kemudian dinilai sesuai. Ketika instansi, masyarakat, dan pemangku kepentingan lainnya berdasarkan alasan tersebut, mereka akan lebih cenderung menggunakan evaluasi untuk perbaikan program.

Faktor yang memperumit, tentu saja, adalah bahwa pemangku kepentingan yang dapat membawa standar yang berbeda dan bahkan bertentangan. Perbedaan nilai dan standar yang diidentifikasi selama pelibatan pemangku kepentingan di Langkah 1. Perspektif pemangku kepentingan tersebut juga akan terlibat dalam deskripsi program dan evaluasi.

7.6 Langkah 6: Diseminasi Hasil Evaluasi

Tujuan akhir dari evaluasi program adalah menggunakan informasi untuk meningkatkan program.

Tujuan yang Anda lakukan di awal proses evaluasi harus memandu penggunaan hasil evaluasi. Hasil evaluasi dapat digunakan untuk menunjukkan efektifitas program Anda, Mengidentifikasi cara untuk meningkatkan program Anda, memodifikasi perencanaan program, menunjukkan akuntabilitas, dan membenarkan pendanaan.

Penggunaan tambahan termasuk yang berikut:

- Untuk menunjukkan kepada legislator atau pemangku kepentingan lainnya bahwa sumber daya dibelanjakan dengan baik dan bahwa program tersebut efektif.
- Untuk membantu dalam membentuk anggaran dan membenarkan alokasi sumber daya.
- Membandingkan hasil dengan tahun-tahun sebelumnya.
- Untuk membandingkan hasil aktual dengan hasil yang diinginkan.
- Untuk menyarankan hasil yang diharapkan realistis.
- Untuk mendukung perencanaan tahunan dan jangka panjang.
- Untuk perhatian pada isu-isu penting untuk program Anda.
- Untuk mempromosikan program Anda.
- Untuk identifikasi mitra untuk kolaborasi.
- Untuk meningkatkan program citra Anda.
- Untuk mempertahankan atau meningkatkan pendanaan.
- Memberikan Arahan bagi program staf.
- Untuk identifikasi kebutuhan pelatihan dan bantuan teknis.

Apa yang terlibat dalam memastikan penggunaan dan berbagi pelajaran? Lima elemen penting dalam memastikan bahwa temuan dari evaluasi yang digunakan:

1. Rekomendasi
2. Persiapan
3. Umpan Balik
4. Tindak lanjut
5. Penyebaran

Rekomendasi adalah tindakan yang perlu dipertimbangkan sebagai hasil evaluasi. Rekomendasi dapat memperkuat evaluasi ketika mereka dan bereaksi

terhadap apa yang ingin diketahui pengguna, dan dapat mengetahui evaluasi jika tidak didukung oleh bukti yang memadai, atau tidak sesuai dengan nilai-nilai pemangku kepentingan. Rekomendasi akan mendukung audiens dan tujuan evaluasi.

Evaluasi adalah alat praktis yang dapat digunakan negara untuk menyampaikan program dan menilai penilaian. Evaluasi program harus diintegrasikan dengan baik ke dalam perencanaan, pelaksanaan, dan pengelolaan program kesehatan masyarakat sehari-hari. Evaluasi program melengkapi prinsip-prinsip operasi CDC untuk kesehatan masyarakat, yang mencakup penggunaan ilmu pengetahuan sebagai dasar pengambilan keputusan dan tindakan, peningkatan keadilan sosial, kinerja efektif sebagai lembaga layanan, dan upaya upaya untuk mencari hasil. Prinsip-prinsip ini perlunya program untuk mengembangkan rencana yang jelas, kemitraan inklusif, dan sistem umpan balik yang dukungan perbaikan berkelanjutan. CDC berkomitmen untuk menyediakan alat tambahan dan bantuan teknis kepada negara bagian dan mitra untuk membangun dan meningkatkan kapasitas mereka untuk evaluasi.

GLOSARIUM

- Akuntabilitas: Tanggung jawab manajer program dan staf untuk memberikan bukti kepada pemangku kepentingan layanan dan dana dari suatu program dan sesuai dengan cakupan, hukum, dan persyaratan fiskal.
- Akurasi: Sejauh mana evaluasi itu benar atau valid dalam apa yang dikatakannya tentang suatu program, proyek, atau materi.
- Aktivitas: Peristiwa atau tindakan aktual yang terjadi sebagai bagian dari program.
- Analisis biaya-manfaat: Analisis yang menggabungkan manfaat program dengan biaya program. Manfaat dan biaya diubah menjadi istilah moneter.
- Analisis biaya-manfaat atau efektivitas biaya *ex ante*: Analisis -manfaat atau efektivitas biaya yang tidak memperkirakan manfaat dan biaya aktual dari suatu program tetapi menggunakan biaya dan manfaat yang dihipotesiskan sebelum fakta. Jenis analisis yang digunakan untuk tujuan perencanaan daripada untuk evaluasi.
- Analisis efektivitas biaya: Analisis yang menggabungkan biaya program dan efek (dampak). Namun, dampak tersebut tidak harus diubah menjadi manfaat atau biaya moneter.
- Analisis statistik deskriptif: Angka dan tabulasi yang digunakan untuk meringkas dan menyajikan informasi kuantitatif secara ringkas.
- Atribusi: Estimasi sejauh mana setiap hasil yang diamati yang disebabkan oleh suatu program, yang berarti bahwa program tersebut telah menghasilkan efek tambahan.
- Bias non-respons: Potensi kemiringan karena non-respons. Jawaban dari unit sampling yang menghasilkan informasi mungkin berbeda pada item yang menarik dari jawaban dari unit sampling yang tidak menjawab.
- Bias penguji: Perubahan yang diamati dalam eksperimen semu yang mungkin merupakan hasil dari keakraban yang berlebihan dengan alat ukur. Ini merupakan ancaman potensial terhadap validitas internal.
- Cost-effectiveness analysis*: Sebuah analisis biaya-manfaat atau efektivitas biaya yang terjadi setelah suatu program beroperasi selama beberapa waktu dan digunakan untuk menilai biaya aktual dan manfaat aktual.
- Data *cross-sectional*: Data yang dikumpulkan pada satu titik waktu dari berbagai entitas.
- Data kualitatif: pengetahuan yang bersifat kategori daripada numerik, sering kali melibatkan, sikap, persepsi, dan niat.
- Data longitudinal: Data yang dikumpulkan selama periode waktu tertentu, terkadang melibatkan aliran data untuk orang atau entitas tertentu dari waktu ke waktu.
- Data primer: Data yang dikumpulkan oleh tim evaluasi khusus untuk studi evaluasi.
- Desain eksperimen semu: Struktur studi yang menggunakan kelompok pembandingan untuk menarik kesimpulan tetapi tidak menggunakan pengacakan untuk membuat kelompok perlakuan dan kontrol. Kelompok perlakuan biasanya diberikan. Kelompok kontrol dipilih untuk kelompok perlakuan sedekat mungkin sehingga kesimpulan tentang dampak tambahan dari program dapat dibuat.
- Desain implisit: Sebuah desain tanpa kelompok kontrol formal dan di mana pengukuran dilakukan setelah paparan program.

Difusi atau imitasi pengobatan: Responden dalam satu kelompok mendapatkan efek yang ditujukan untuk kelompok (program) pengobatan. Ini merupakan ancaman bagi validitas internal.

Efek interaksi: Efek bersih gabungan dari dua (atau lebih) variabel yang mempengaruhi hasil eksperimen semu.

Evaluasi: Pengumpulan informasi yang sistematis untuk mendokumentasikan dan menilai suatu program yang diimplementasikan dan dioperasikan.

Evaluasi desain: Model logistik atau kerangka konseptual yang digunakan untuk sampai pada kesimpulan tentang hasil.

Evaluasi program: Pengumpulan informasi yang sistematis tentang kegiatan, karakteristik, dan hasil program untuk menilai tentang program, meningkatkan efektivitas program, dan/atau informasi keputusan tentang pengembangan program di masa depan.

Evaluasi strategi: Metode yang digunakan untuk mengumpulkan bukti tentang satu atau lebih hasil dari suatu program. Strategi evaluasi terdiri dari desain evaluasi, metode pengumpulan data, dan analisis teknik.

Inferensi kausal: Proses logistik yang digunakan untuk menarik kesimpulan dari bukti mengenai apa yang telah dihasilkan atau "disebabkan" oleh suatu program. Mengatakan bahwa suatu program menghasilkan atau menyebabkan hasil tertentu berarti bahwa, jika program itu tidak ada (atau jika telah ada dalam bentuk atau derajat yang berbeda), maka hasil yang diamati (atau tingkat hasil) tidak akan terjadi.

Instrumentasi: Pengaruh perubahan alat ukur dari satu pengukuran ke pengukuran lainnya, seperti ketika wawancara yang digunakan berbeda. Ini merupakan ancaman bagi validitas internal.

Keandalan: Sejauh mana pengukuran, ketika berulang kali diterapkan pada situasi tertentu secara konsisten menghasilkan hasil yang sama jika situasi tidak berubah di antara aplikasi.

Kebijakan moneter: Tindakan pemerintah yang mempengaruhi jumlah uang beredar dan tingkat suku bunga. Bisa juga dalam bentuk program.

Kedalaman: Tingkat akurasi dan detail pengukuran.

Kelengkapan : Cakupan yang luas dan mendalam tentang evaluasi isu-isu yang menarik.

Kelompok fokus: Sekelompok orang yang dipilih karena relevansinya dengan evaluasi yang dilibatkan oleh para pengamat dalam diskusi yang dirancang untuk wawasan, ide, dan pengamatan tentang topik yang menjadi perhatian.

Kelompok kontrol: Dalam desain eksperimen semu, sekelompok subjek yang menerima semua pengaruh kecuali program dengan cara yang bertahan sama dengan kelompok perlakuan (yang terakhir disebut, dalam beberapa keadaan, kelompok eksperimen atau program). Juga disebut sebagai kelompok nonprogram.

Kelompok pembanding: Kelompok yang tidak terkena program atau pengobatan. Juga disebut sebagai kelompok kontrol.

Kelompok perlakuan: Dalam desain penelitian, kelompok mata pelajaran yang menerima program. Juga disebut sebagai kelompok eksperimen atau program.

Keluasan : Cakupan Cakupan Pengukuran.

Kepatutan: Sejauh mana evaluasi telah dilakukan dengan cara yang membuktikan

- kepatuhan tanpa kompromi terhadap prinsip dan cita-cita tertinggi (termasuk etika profesional, hukum perdata, kode moral, dan perjanjian kontrak).
- Metode analitik langsung: Metode yang digunakan untuk memproses data untuk memberikan bukti tentang dampak langsung atau hasil dari suatu program.
- Metode pengumpulan data: Cara mengumpulkan fakta tentang suatu program dan hasilnya. Metode pengumpulan data yang sering digunakan dalam program evaluasi meliputi literatur pencarian, file, observasi alam, survei, pendapat ahli, dan studi kasus.
- Model logika: Cara sistematis dan visual untuk menyajikan hubungan yang dirasakan antara sumber daya yang Anda miliki untuk mengoperasikan program, kegiatan yang Anda rencanakan, dan perubahan atau hasil yang ingin Anda capai.
- Model ekonomi makro: Sebuah model interaksi antara barang, tenaga kerja, dan pasar aset ekonomi. Model ini berkaitan dengan tingkat output dan harga berdasarkan interaksi antara permintaan dan penawaran agregat.
- Model ekonomi mikro: Sebuah model perilaku ekonomi pembeli dan penjual individu, di pasar tertentu dan keadaan keadaan.
- Non-respons: Situasi di mana informasi dari unit sampling tidak tersedia.
- Pemangku kepentingan: Orang atau organisasi yang diinvestasikan dalam program atau yang tertarik dengan hasil evaluasi atau apa yang akan dilakukan dengan hasil evaluasi.
- Rencana evaluasi: Sebuah dokumen tertulis yang menjelaskan pendekatan atau desain yang akan digunakan untuk memandu evaluasi. Ini mencakup apa yang akan dilakukan, bagaimana akan dilakukan, siapa yang akan melakukannya, kapan akan dilakukan, evaluasi dilakukan, dan bagaimana temuan kemungkinan akan digunakan.
- Sampling probabilitas: Pemilihan unit dari populasi berdasarkan prinsip pengacakan. Setiap unit populasi memiliki probabilitas terpilih (bukan nol) yang dapat dihitung.
- Standar: Prinsip yang umumnya disepakati oleh para ahli dalam pelaksanaan dan evaluasi untuk ukuran atau kualitas evaluasi (misalnya, ketepatan, kegunaan, kegunaan).
- Studi kasus: Metode pengumpulan data yang melibatkan studi tentang kasus atau proyek tertentu dalam suatu program. Metode itu sendiri terdiri dari satu atau lebih metode pengumpulan data (seperti wawancara dan file).
- Survei: Metode pengumpulan data yang melibatkan upaya terencana untuk mengumpulkan data yang diperlukan dari sampel (atau sensus lengkap) dari populasi yang relevan. Populasi yang relevan terdiri dari orang atau entitas yang terkena program dampak (atau orang atau entitas serupa).
- Tingkat kepercayaan: Pernyataan bahwa nilai sebenarnya dari suatu parameter untuk suatu populasi berada dalam kisaran tertentu dengan tingkat probabilitas tertentu.
- Utilitas: Sejauh mana evaluasi menghasilkan dan menyebarkan laporan audiens yang relevan dan memiliki dampak pada pekerjaan mereka.
- Validitas eksternal: Kemampuan untuk menggeneralisasi kesimpulan tentang suatu program ke masa depan atau kondisi yang berbeda. Ancaman terhadap validitas eksternal meliputi seleksi dan interaksi program, setting dan program interaksi, serta sejarah dan interaksi program.
- Validitas kesimpulan: Kemampuan untuk menggeneralisasi kesimpulan tentang program yang ada ke tempat, waktu, atau situasi lain. Kedua masalah validitas internal dan

eksternal harus ditangani jika kesimpulan tersebut ingin dicapai.

Validitas internal: Kemampuan untuk menyatakan bahwa suatu program telah menyebabkan hasil yang dicapai (sampai tingkat tertentu), di hadapan penjelasan alternatif yang masuk akal. Ancaman yang paling umum untuk validitas internal adalah sejarah, pematangan, kematian, bias seleksi, artefak regresi, difusi, dan imitasi pengobatan dan pengujian.

Validitas pengukuran: Suatu pengukuran valid sejauh itu mewakili apa yang dianggap dan dianggap mewakili. Langkah-langkah yang valid tidak memiliki bias sistematis.

REFERENSI

HM Treasury. (2018). The Green Book: Central Government Guidance on Appraisal and Evaluation. [pdf] London. Crown Copyright. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/685903/ 2018.

UK Research and Innovation. Medical Research Council. **Process evaluation of complex interventions**. 2015

United Nations Development Programme. **UNDP Handbook on Planning, Monitoring and Evaluating for Development Results**. 2006

U.S. Department of Health and Human Services Centers for Disease Control and Prevention Office of the Director, Office of Strategy and Innovation. **Introduction to program evaluation for public health programs: A self-study guide**. Atlanta, GA: Centers for Disease Control and Prevention, 2011.

World Health Organization. **Evaluation in health promotion: principles and perspective**. 2001 Available at <https://www.who.int/publications/i/item/9789289013598>

ISBN 978-623-5481-80-7 (PDF)



9 786235 481807



Penerbit
CV. PATRA MEDIA GRAFINDO
BANDUNG

Jl. Sekeloa Tengah No. 100, Bandung
Telp. (022) 25011111, 25011112
Email: info@patramedia.com
Website: www.patramedia.com