Canadian Academy of Health Sciences
ASSESSMENT SUMMARY

MAKING AN IMPACT

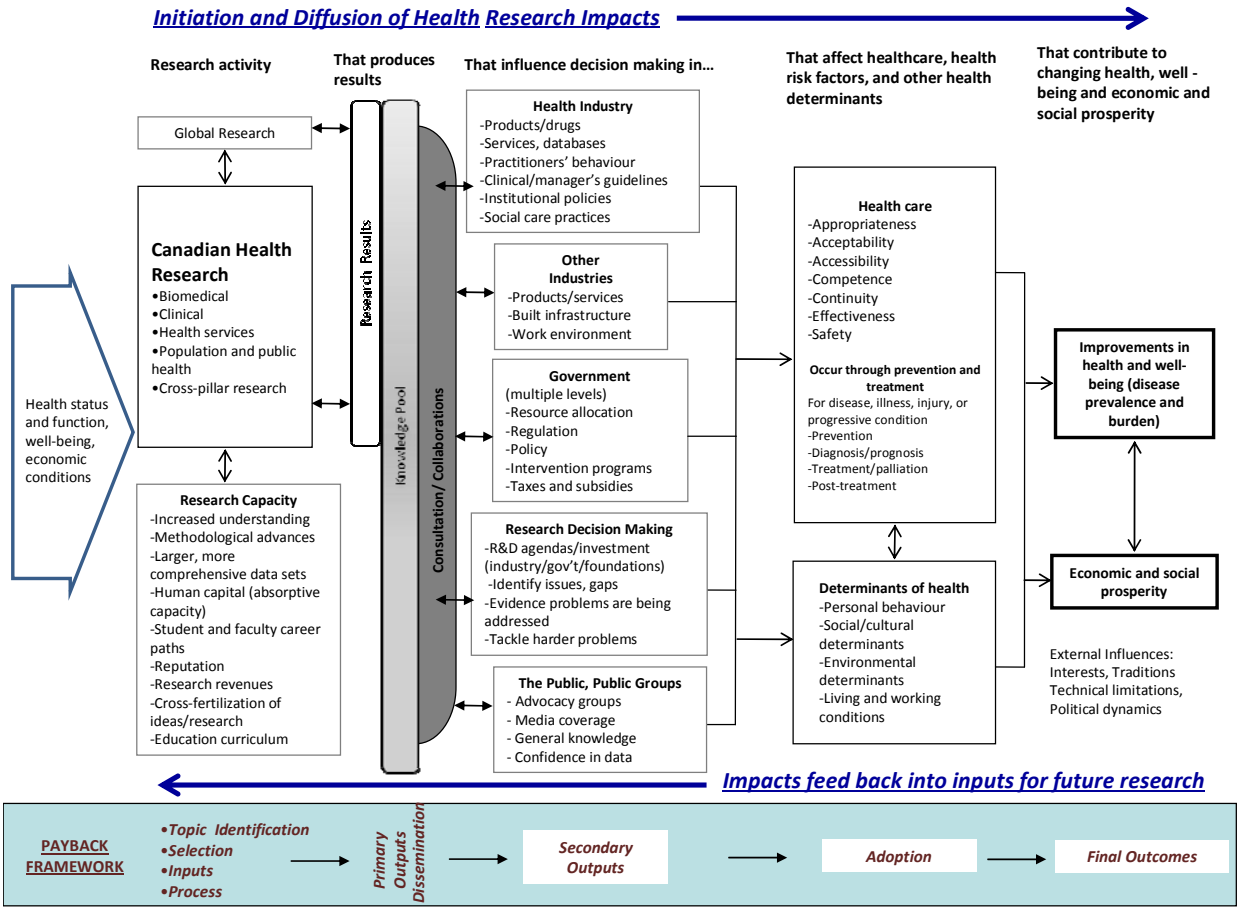
A Preferred Framework and Indicators to Measure Returns on Investment in Health Research

Summary of the Report on the Return on Investments in Health Research
January 2009

Executive Summary

Twenty-three different organizations sponsored this assessment. They all share an interest in defining the impacts of health research and learning how to improve the returns on investments in health research. Our remit from these sponsors was: Is there a "best way" (best method) to evaluate the impacts of health research in Canada, and are there "best metrics" that could be used to assess those impacts (or improve them)?

Based on our assessment, we propose a new impacts framework and a preferred menu of indicators and metrics that can be used for evaluating the returns on investment in health research.



The CAHS impact framework demonstrates how research activity informs decision making, eventually resulting in changes in health and economic and social prosperity (left to right arrow). The framework also shows how research impacts feed back upstream, potentially influencing the diffusion and impacts of other research, and creating inputs for future research (right to left arrow).

This framework builds on the combined logic model and impacts approach of the "payback model" (Buxton, M.J., and Hanney, S.R., 1996 – adapted by CIHR in Canada in 2005 and 2008), revised by our panel into a "systems approach" to capture impacts (this is shown at the bottom of the Figure). It is

designed to be used as a roadmap to track health-research impacts in five main categories: 1) advancing knowledge, 2) building capacity, 3) informing decision-making, 4) health impacts, and 5) broad socio-economic impacts.

Each of the main categories consists of subcategories, which identify and partition evaluation methods, and data that permit impact evaluation and identification of contributing factors. The first-level subcategories of each impact category are as follows:

- *Advancing knowledge indicators and metrics* include measures of research quality, activity, outreach and structure. We have also identified some aspirational indicators of knowledge impacts using data that are highly desirable but currently difficult to collect and/or analyze (such as an expanded relative-citation impact that covers a greater range of publications, including book-to-book citations and relative download-rates per publication compared to a discipline benchmark).
- *Research capacity-building indicators and metrics* fall into subgroups that represent personnel (including aspirational indicators for improving receptor and absorptive capacity), additional research-activity funding and infrastructure.
- *Informing decision-making indicators and metrics* represent the pathways from research to its outcomes in health, wealth and well-being. They fall into health-related decision-making (where health is broadly defined to include health care, public health, social care, and other health-related decisions such as environmental health); research decision-making (how future health research is directed); health-products industry decision-making; and general public decision-making. We also provide two aspirational indicators for this category (media citation analysis and citation in public policy documents).
- *Health-impact indicators and metrics* include those on health status, determinants of health and health-system changes, and they include quality of life as an important component of improved health. Determinants of health indicators can be further classified into three major subcategories: modifiable risk factors, environmental determinants, and modifiable social determinants.
- *Broad economic and social impacts* are classified into activity, commercialization, health benefit (specific costs of implementing research findings in the broad health system), well-being, and social-benefit indicators (socio-economic benefits).[2]

By choosing appropriate sets of indicators from the menu presented below, the CAHS impacts framework can be used to track impacts within any of Canada's four "pillars of health research" (basic biomedical, applied clinical, health services and systems, or population health) or within domains that cut across these pillars. It can also be used to describe impacts at various levels — individual,

[2] The indicators in these last two categories represent changes in impacts, which may or may not be due to research findings. Without tracking these changes, however, identifying the impacts of research would be impossible.

institutional, provincial, national or international — and to define funders' "returns" by (eventually) quantifying the value(s) of impacts to end-users as a function of dollars invested.

We believe that the framework's breadth, depth and flexibility fulfills our sponsors' request that the criteria be:

- useful to a full range of funders/research types,
- compatible with what is already in place in Canada,
- transferable to international comparisons, and
- able to identify the full spectrum of potential impacts.

To explain the logic behind the framework diagram, we must begin on the left, recognizing that health research occurs within the context of the Canadian population's health status, functioning, well-being and economic conditions. Canadian research is conducted within the context of a much larger volume of global health research. The Canadian context is further defined by the people and infrastructure that comprise current national research capacity. This influences selection of research questions within the four main pillars of Canadian health research, and domains that cut across these pillars. Moving to the right, the results of funded health research are first evident in published research documents, which then add to the global pool of knowledge. To varying degrees, the public and private industrial sectors (health and otherwise), as well as governments, draw upon this growing pool of knowledge to inform and influence future research agendas. New knowledge is also more broadly disseminated, usually by the media, to the general public and various interest groups.

It is important to note that these more direct uses of the products of health research are not their only impact. The end-products of health research manifest themselves partly in better and more appropriate interventions by the health system (both treatments and prevention), in health-improving changes to the broader determinants of health and, finally, in a more prosperous economy and society that is the indirect result of a healthier Canadian population.

While the CAHS panel clearly recognizes that making all the linkages from pieces of funded research at the left of the diagram to the outcomes at the right is a major challenge, with many gaps to be filled, it is precisely these outcomes that are the foundation for the public's support of health research. Moreover, we must emphasize that population health and economic prosperity are influenced by many factors other than health research. Thus a major question remains about how to parcel out the contributions of health-research to these bottom-line outcomes. This "attribution problem" for health research impacts is the main reason that the panel recommends that health-research funders support research into the basic science of health-research impacts.

To apply the framework appropriately, it should be populated by sets of both metrics and indicators of impact that are chosen to address specific evaluation questions. Users should take the following steps:

1. Define and prioritize *specific evaluation question(s).*
2. Use the framework to determine where to look for impacts.
3. Choose the impact categories (and subcategories) of interest: advancing knowledge, capacity building, informing decision-making, health impacts, and broad economic and social impacts. Be as specific as possible about where impacts are expected to occur and at what level (see subcategories above).

4. Choose a set of indicators and metrics carefully from the appropriate categories and subcategories of interest that will address the evaluation questions.
5. Ensure that sets of indicators meet the criteria of attractiveness and feasibility noted below.
6. Avoid inappropriate uses of the framework: Users should not ignore potential undesirable impacts of research by seeking only positive impacts (the so-called "Halo Effect"). They should also avoid "double-counting" of research impacts, and they should not ignore the "attribution issue" (in which contributions of other research and other factors known to impact health outcomes are ignored). Finally, users should consider "the counterfactual" — what would have happened if the research was not performed?. These issues are not easily resolved, and each requires additional research, including development of new indicators and metrics. Until that research is done, health-research impact evaluations should be based on balanced evidence and remain cautious about attribution. National and international collaboration in resolving these issues is highly recommended.

Our assessment also suggests that it is not possible to prescribe sets of indicators for different sponsoring groups (such as universities, government funders, voluntary organizations, etc). Users must define their own questions carefully and choose sets of metrics and indicators to address those questions in an objective way. In our report (on pages 106-108), we give three examples of how to use the framework to address theoretical evaluation questions for three potential users with quite different needs:

a. the biotechnology cluster (which is interested in the quality of its research, evidence of commercial gain, or whether employment opportunities were created);
b. a provincial funder of project grants (who is mainly interested in research quality, evidence of any health gains, and the ability to attract additional research investments to the province); and
c. a federal funder of research fellowships (who is interested in research quality, increased skill sets in Canada and dissemination of knowledge by its fellows).

These examples hint at the diversity of potential evaluation questions and why we cannot prescribe simple sets of metrics for all. Evaluation is expensive, and its costs must be reconciled by the potential magnitude of the findings' impact on users (funders and other stakeholders). Thus, we recommend collaboration in establishing evaluation questions of national interest in order to achieve economies of scale.

The second part of our mandate was to define appropriate metrics of health-research impact and to describe how to use them. We began this exercise by trying to define what precisely constitutes a 'metric', and by identifying how many such metrics exist. For our purposes, the term 'metric of impact' was taken to mean 'a numerical measure of impact'. Our assessment suggested that, while some metrics of impact (by that definition) clearly exist, many other useful non-numerical 'indicators of impact' have been described. The term, "indicators of impact", was therefore adopted to define all

indicators — some of which are numerical (metrics), and many of which are descriptive or qualitative. We have thus defined a menu of indicators that includes metrics as a subset.

Before listing the indicators, we will define how to use them. No single indicator of impact of health research — whether in any domain or in any impact category — is sufficient to demonstrate impact for any organization. Any single indicator can be dismissed as being partial, imperfect and likely to distort. For example, using publication counts as a single indicator is "perverse" and could have detrimental effects on scientists by sending the message that "volume equals quality." However — as part of a set of indicators that includes quality-of-research indicators, funding-received indicators (capacity building) and indicators for translating research findings for a variety of stakeholders — examining publication counts can help to identify those researchers producing quality *and* volume of research outputs. Indicator sets must also be appropriate to the right "level of aggregation" within each impact category. Some are appropriate at the individual researcher level, while others are not.

Selection of sets of indicators and metrics should be:

- focussed on the objectives of the organizations that will use them,
- appropriate for the stakeholders who are likely to use the information,
- balanced to cover all significant areas of work performed by an organization,
- robust enough to cope with organizational changes (such as staff changes),
- integrated into management processes, and
- cost-effective (balancing the benefits of the information against the costs of collection).

Indicators should also meet criteria of attractiveness and feasibility listed in the table below:

Table of preferred indicator characteristics - While it is not possible for every indicator to meet all of these criteria, it is important that each *set of indicators and metrics be selected to do so*.

Attractiveness:

Validity – does the indicator or metric reasonably reflect the underlying concept or construct that it is intended to measure?

Relevance – does the indicator or metric relate directly to a critical aspect of the research?

Behavioural impact – does the indicator or metric drive behaviour in a particular direction? Is it likely to result in any negative, unintended consequences? Does it create "perverse incentives?"

Transparency – is the methodology, and the strengths and weaknesses relating to the indicator or metric, readily apparent?

Coverage – does the indicator or metric cover a large proportion of output from research to be assessed?

Recency – do the data relate to current research performance, or look over a longer timescale?

Methodological soundness – is the calculation of the metric sound and statistically robust?

Replicability – can others reproduce the indicator or metric, and can it be used year on year in a comparable fashion?

Comparability – do other organizations collect comparable information or have targets to benchmark against?

Feasibility:

Data availability - do the data required to derive indicators or metrics exist, and do both the analysts and those being assessed have access to it?

Cost of data – how expensive is it to purchase the data outright or obtain on license?

Compliance costs – how labour intensive is it to extract/obtain the data?

Timeliness – can the data be obtained/provided relatively quickly?

Attribution – can the data be discretely ascribed to the unit being assessed? Direct attribution is ideal, but unlikely given current knowledge and methods; using attribution as a concept is important, as it provides a link between the impact seen and the research.

Avoids gamesmanship – does the indicator or metric provide scope for special interest groups or individuals to game the system?

Interpretation – can the data be open to misinterpretation or misuse by commentators and/or actors using the evaluation findings?

Well-defined – does the metric have a clear, unambiguous definition so that data will be collected consistently, and so that the measure is easy to understand and use?

Our menu of indicators and metrics is shown below, arranged according to the impact categories in the framework: advancing knowledge, capacity building, informing decision-making, health impacts, and broad economic and social impacts. These are the most wide-ranging indicators available, and they aim to capture as much information as possible across all health research. For each indicator and metric in the table, we have identified the "level" at which it is most appropriately applied (for example, individual researchers, research groups, institutions, etc.) and the CIHR pillar(s) of research to which each relates. Many are validated and practical, as they use existing data. However, not all of these indicators and metrics are currently available "off the shelf". In some categories we suggest indicators that represent important constructs, but whose characteristics have not yet been fully developed – we call these "aspirational indicators". It is necessary to prioritize and invest in developing and collecting appropriate data for all of these indicators, to prioritize aspirational indicators, and add to the list of indicators. It is also important to note that all but the most proximal indicators in the framework suffer from the issue of attribution (they are likely influenced by factors in addition to health research) and they require the most future development.

The quality of the indicator data is paramount. Special attention should be given to collecting data in a standardized fashion to allow comparisons of research impacts across funders within Canada and, ideally, outside of Canada. Standardizing the definitions for disease groups and types of research is the first step (such as the UKCRC classification (UK Clinical Research Collaboration 2007)).

Unfortunately, our assessment suggests that the data already being collected in Canada are not sufficient to provide all the information an evaluator might want, as there are likely significant gaps with regard to secondary outputs and outcomes. To address these gaps, there are two data collection techniques in use elsewhere: end-of-grant reporting, and common CVs for researchers. The standard format for end-of-grant reporting is a qualitative description of the research findings, process, and likely outcomes. Recent work in the U.K. incorporates a payback model-based survey that allows researchers to report on outputs and outcomes quickly at the end of grants, and again five years after the grant—thereby identifying late-arising impacts (Wooding 2008). It is also possible to expand the common CV (The Common CV System 2006) to include a variety of standardized outputs from researchers, such as presentations to public audiences, or consultations to government. Using these two approaches for collecting data ensures that there will be opportunities to identify research diffusion from researchers to a variety of stakeholders without being too onerous.

CAHS Menu of Preferred Indicators and Metrics of Impact - Appropriate indicators and metrics are arranged according to CAHS framework impact category. They should be selected in sets and mapped onto the CAHS framework to address different evaluation questions. They were selected for this menu from over 300 current indicators that were considered by the panelists. The subset of numerical indicators is called 'metrics' of impact. Note: this table contains our "starting menu" of preferred indicators and metrics - we recommend that they should be expanded over time.

ADVANCING KNOWLEDGE					
Category	**Indicator**	**Description**	**Level of Application**	**Comments**	**Pillars that indicators are relevant to**
QUALITY	Relative citation impact	* Average citations received by the unit being analyzed, compared to the world citation rate for the discipline(s) * World citation rates per discipline should be made widely available to interested parties	* Individual - not recommended * Group/department/grant - recommended * Institution/funding agency - recommended * National - recommended	* Must use discipline-specific benchmarks to account for different citation practices across disciplines * Only robust if based on a sufficient set of publications (individual researchers generally produce too few for robust analysis)	All pillars
	Highly cited publications	Individual publications are assessed against world citation thresholds to determine if they are in the top 1%, 10%, etc. of most highly cited publications in the world in that research area	Recommended at all levels	Must use discipline-specific benchmarks to account for different citation practices across disciplines	All pillars
	Publications in high-quality outlets (or desired outlets)	* Proportion of publications (publishers, conferences, journals) that appear in outlets judged to be of high quality * Could also include outlets that target specific stakeholders, such as those used by health practitioners	Recommended at all levels	Activity in a number of countries where disciplines are engaged in ranking the outlets of their discipline, including ranking publishers as well as journals (e.g. ESF Humanities project; Australia ERA journal and publisher rankings)	All pillars, but likely to be more important for pillars III and IV, where a smaller proportion of knowledge production is in journals
ACTIVITY	Share of publications	Number of publications from the unit under study as a proportion of a reference output (usually the level of aggregation above the unit under study)	* Individual - not recommended * Group/department/grant - recommended (share of institutional/funding body output) * Institution/funding agency - recommended (share of national output) * National - recommended (share of world publications)	* Normally done for field of research, rather than total publications * Can currently be easily calculated for indexed journal articles, but not for other types of publications	All pillars
	Publication counts	* Simple counting of outputs * Can be useful for new researchers who have no publication record allowing citation analysis	* Individual - recommended (number of publications by type: journal articles, books, book chapters, conferences, etc.) * Group/department/grant - not recommended * Institution/funding agency - not recommended * National - not recommended	* Counts by themselves are a poor indicator * The data are routinely collected in order to calculate other indicators (e.g. publication share, relative citation impact) * There needs to be a comparative aspect (e.g. is the level of output above or below that expected in that discipline)	This indicator is more important in pillar III and IV where a smaller proportion of knowledge production is in the journal literature. We strongly recommend that this indicator not be used as an indicator of quality in pillars I or II

ADVANCING KNOWLEDGE					
Category	**Indicator**	**Description**	**Level of Application**	**Comments**	**Pillars that indicators are relevant to**
OUTREACH	Co-author analysis	Determining the proportion of publications that are co-authored internationally, nationally, with industry, with other disciplines, etc.	Recommended at all levels	The selection of type of co-authorships to be analyzed will depend on the focus of the analysis	All pillars
	Field analysis of citations	Determining the proportion of citations that come from articles in the same field, and which other fields	Recommended at all levels	Gives an indication of the interdisciplinarity of the research by demonstrating the pick-up of research outside the core discipline	All pillars
CONTEXTUAL / STRUCTURAL	Relative activity index	* Determining the fields of research in which a unit is most strongly focussed * Uses the number of HCPs in each research area to show activity that is highest quality only	* Individual - not recommended * Group/department/grant - not recommended * Institution/funding agency - recommended * National - recommended	The benchmark for assessment will vary according to the research question, e.g. an institution may wish to compare its output to the national distribution, while at the national level the comparison might be to the world distribution or to similar countries	All pillars
ASPIRATIONAL INDICATORS	Expanded relative citation impact	Expanding citation analysis to cover a greater range of publications, including book-to-book citations	Aspirational at all levels except for the individual	There is work going on to try to improve the citation databases to include additional resources such as books, and this could be in place in the near future	Could prove especially important for pillars III and IV where a greater proportion of output is in the non-journal literature
	Relative download rate	Average number of downloads per publication compared to discipline benchmark	Aspirational at all levels except for the individual	* Ideally, downloads should differentiate between audiences, i.e. downloads from academic institutions, government agencies, general public, etc. * An equivalent indicator to highly cited publications for individuals could be "most downloaded"	Could prove especially important for pillars III and IV where a greater proportion of output is in the non-journal literature
	Research diffusion	Based on end-of-grant reports, which should include named individual researchers who should benefit from the research, and a sample of such individuals and their assessments of the actual usefulness of the research results, qualitative assessment of diffusion / uptake of research results	Aspirational at all levels	Requires thorough end-of-grant reports and follow-up	All pillars

CAPACITY BUILDING					
Category	**Indicator**	**Description**	**Level of Application**	**Comments**	**Pillars that indicators are relevant to**
PERSONNEL	Graduated research students in health- related subjects	* Numbers of graduated PhD/MSc/MDs, year on year * Should be able to disaggregate to subjects, gender, etc.	* Not recommended at the individual level * Can be used at institutional level * Most useful provincially or nationally	* As an aspiration we would also like to track the success of training programs in producing outstanding scientists and the progress that all research graduates make * Could be done in part using the Statistics Canada National Graduate Survey	All pillars
	Numbers of research and research- related staff in Canada	* Split into researchers, research assistants, and other staff * Can be disaggregated by province, research sector, etc.	* Not recommended at the individual level * Can be used at institutional level * Most useful provincially or nationally	Data already collected by Statistics Canada	All pillars
FUNDING	Levels of additional research funding	Funding from "external" sources that can be attributed to the capacity built in an organization, institution, or region. Could also include matched funding	Only recommended for funders, provinces, and nationally	Difficult to attribute to research funded by that province/organization, since researchers tend to be funded by multiple funding bodies (risks double counting)	All pillars
INFRASTRUCTURE	Infrastructure grants ($)	The amount in dollars of infrastructure funding pulled in by a research project, group, organization	Only recommended for institutions, organizations, provincially, and nationally	Captures the different aspects of infrastructure (kit, databases, buildings) since they all come from infrastructure grants, but misses out on infrastructure from other sources (e.g. university re-allocation of space, etc.) NOTE: This can be perverse if not aligned with operating money	All pillars
	% of activity grants with infrastructure support	Co-ordination of infrastructure grants with activity grants by identifying which activity grants have received additional infrastructure support to allow the research to occur	Only recommended for institutions, organizations, provincially, and nationally	* Does not account for research that has no new infrastructure costs or ones that are covered by universities * Data collection may be difficult and may have to be through surveying activity grant holders	All pillars
ASPIRATIONAL INDICATORS	Receptor capacity	Ability of those in policy and administrative positions to take research findings on board	Unlikely to be able to link to specific research findings, but could track the development of receptor capacity in Canada	There are surveys available to test receptor capacity, although these tend to be associated with specific training schemes	All pillars, particularly III and IV
	Absorptive capacity	Ability of researchers to take on other research from outside their organization, country, etc. and exploit that knowledge	Could address absorptive capacity for organizations, provinces, or nationally	Most commonly attributed through collaborations (particularly industry - academia collaborations) or R&D funding intensity	All pillars

INFORMING DECISION MAKING						
Category	**Subcat-egory**	**Indicator**	**Description**	**Level of Application**	**Comments**	**Pillars that indicators are relevant to**
HEALTH RELATED[3]	***Health care***	Use of research in guidelines	Analyzing citations to research in clinical and service guidelines	* Can be applied for individual researchers * More practical at aggregate levels (group/institution/ province/nation)	Allows identification of specific research informing health care and proportion of Canadian research informing health care	Mainly pillars I, II, and III
	Public health	Survey of public health policy makers	Asking public health policy makers what research has been used to inform their policies.	* Unlikely to be useful for individuals * May be useful for groups * Dependent on level of detail provided by policy makers	Surveying may be difficult unless policy makers are incentivized to take part	Likely to be pillars III and IV
	Social care	Use of research in guidelines	Analyzing citations to research in social care service guidelines	* Can be applied for individual researchers * More practical at aggregate levels (group/institution/ province/nation)	Allows identification of specific research informing social care and proportion of Canadian research informing social care	Likely to be pillars II, III, and IV
	Other	Researcher reported use of findings outside health	Example: health research findings could be picked up by transport or employment policy to improve safety or working conditions	* Could be applied to individuals * Better used at institution/funder levels	Since there are many different areas within "other," no single top-down indicator can collect all impacts, however, researchers may not know if their research is used outside their area of research	All pillars
	Health-related education	Research cited in ongoing health professional education material	Continuing health professional education materials produced cite research to support new practices	* Can be linked to individuals but likely to be small numbers * More appropriate at group/institution/funder levels * Recommended at provincial and national levels	* There may be issues accessing the references for these materials * Early health professional education covered in "research education"	All pillars
RESEARCH	***Research funding***	Citation analysis of successful funding applications	Identifying cited research in successful funding applications to identify underpinning research informing new research direction	* Can be used for groups and larger aggregations * Not recommended for individuals since number of citations is likely to be small	* Accessing references in successful applications can only be performed by research funders themselves * Data would have to be shared between funders	All pillars
	Research policy	Consulting to policy	Number of consultations to policy makers (from organizational to national policy) by researchers - year-on-year analysis	* Recommended for individuals; can help to identify which individuals are strongly linked into policy circles * Can be aggregated to groups above, but since there is no desired level of consultation is less useful at higher aggregations	* Needs to be addressed through surveying researchers * Top-down approach will miss "un-official" consultation	All pillars
		Requests for research to support policy	Number of requests for research for policy makers; primarily systematic reviews	* Only relevant at a provincial or national scale * Determines the level of interest in research, therefore not something research funders can influence directly	Can be addressed through official requests for research (systematic reviews commissioned) or through researchers' responses to requests	All pillars

[3] Within the four subcategories that represent the different aspects of a broad health system (health care, public health, social care, and other health related systems), there is a three-layer hierarchy of data sources for informing decision making metrics. The top level involves published evidence that identifies research; the middle level, surveying decision makers to identify what has influenced them; and the bottom level, asking researchers to report on how their research has informed decisions. The "most appropriate" indicators identified here are based on the most likely available information for each aspect of health-related decision making (so, if higher levels of information are not readily available, we recommend collecting information at the level below).

INFORMING DECISION MAKING						
Category	**Subcat-egory**	**Indicator**	**Description**	**Level of Application**	**Comments**	**Pillars that indicators are relevant to**
	Research Education	Research used in curricula for new researchers	Citation of research in textbooks and reading lists for university students in health-related disciplines	* Not recommended for individuals * Most useful at group/institution/funder/ province/national levels	Reliant on accessing lists of textbooks and papers used in teaching, as well as mining citation data from them	All pillars
HEALTH PRODUCTS INDUSTRY	***n/a***	Number of patents licensed	* Counts of licensed patents * Can be benchmarked against previous years or against internationally held patents	* Can be used for individuals * Most useful at group/institution/province and national level where sample sizes are larger	Data already maintained on patents licensed in Canada and reported on by Treasury Board	Likely to be pillars I and II
		Clustering/ co-location	Co-location analysis to show where industry is located in relation to academic centres	Only useful at provincial and national levels	Can provide an overview of where innovation and knowledge transfer is likely to occur	Likely to be pillars I and II
		Consulting to industry	Number of researchers consulted by industry; year-on-year values	* Can be used for individuals to identify those translating to industry * For group/institution/provincial levels can show environments conducive to knowledge translation (KT)	Data can be gathered through company reports or through researchers (as part of expanded CV or end of grant reporting)	Likely to be pillars I and II
		Collaboration with industry	Co-author analysis (bibliometric) of collaboration between industry and academia	* Not recommended for individuals (sample size too small) * Recommended for groups/institutions/ provinces/nationally	Reliant on industry publishing research findings in journals	Likely to be pillars I and II
		Use of research in stage reports by industry	Citation analysis of stage reports in development of products by industry	* Not recommended for individuals (sample size) * Recommended for groups/institutions/ provinces/nationally	Relies on accessing stage reports for industry (should be publicly accessible) and the ability to mine citations from them	Likely to be pillars I and II
GENERAL PUBLIC	***Advocacy groups***	Research cited in advocacy publications	Research mentions in publications (leaflets etc.) produced by advocacy groups, including patient organizations	* Not recommended for individuals (sample size) * Recommended for groups/institutions/ provinces/nationally	Misses other work for advocacy groups that is not cited, but consultations for advocacy can be captured in an expanded CV	All pillars
	Public education	Public lectures given	Number of lectures given to public audiences	Individual levels and above	Data could be collected through an expanded standard CV or through end-of-grant reporting	All pillars
ASPIRATIONAL INDICATORS	***Media***	Media citation analysis	Analyzing mentions of research in newspapers	Recommended at the individual level and aggregations above since media tends to mention individuals	* Potential international database of major national newspapers being developed * Requires individuals to identify research mentions in newspapers on a daily basis	All pillars
	Public policy use	Citations in public policy documents	Analyzing citations to research in public policy documents (grey literature)	* Could be applied at the individual level or above * More useful at the group level and above	The advent of Google Scholar as an analysis tool that can access citations in grey literature may help to analyze research informing policy decisions	All pillars

Indicators and metrics in the above sections of the table have a direct link to research. In the *health impacts* and *broad economic and social impacts* tables below, where links to research findings are

much harder to identify, we list the information that is most important to capture to identify changes in health, wealth, well-being, and social circumstances. It is necessary to perform additional studies to determine the link between research and the indicators below.

HEALTH IMPACTS						
Category	**Subcategory**	**Indicator**	**Description**	**Level of Application**	**Comments**	**Pillars that indicators are relevant to**
HEALTH STATUS	*Morbidity to include functional impacts*	Prevalence	Number of cases for a condition in a population (shown as a percentage)	Population level (from subgroups to full population)	Useful to show the impact of a condition on a population	Applicable to all pillars
		Incidence	Number of new cases for a condition per 100,000 population	Population level (from subgroups to full population)	Useful for identifying the new cases of a condition	Applicable to all pillars
	Mortality	PYLL	* Potential Years Life Lost * Number of years of life lost due to premature death (before 75)	Population level (from subgroups to full population)	Already collected across Canada through CIHI and Statistics Canada	Applicable to all pillars
	Quality-adjusted mortality	QALYs	* Quality-adjusted Life Years * Provides a value between 1 (perfect health) and 0 (death) of quality of life for each year lived after an intervention	Can be applied to specific interventions provided that data are collected, and can be used to describe populations	* Useful for linking to research impact since QALYs are linked to interventions (which can be more easily traced to research findings) * At the population level, data source is Canadian Community Health Survey	Applicable to all pillars
		PROMs	* Patient-reported Outcome Measures * Using a standardized questionnaire to determine patient views on quality of care and quality of life pre and post-treatment	Individual patients for clinical practice, but aggregations (e.g. hospital; disease state) for evaluation of research impacts	* Being developed to be more widely used in the UK NHS * Relies on patient reporting of their well-being	Applicable to all pillars
DETERMINANTS OF HEALTH	*Modifiable risk factors*	Example: obesity; alcohol consumption	Measures of prevalence of specific factors; e.g. for obesity, prevalence of BMI>30 for different population groups	* Can be at individual level * More useful for populations or sub-populations	Must be specific for the health problem under investigation	All pillars, but mainly pillar IV
	Social determinants	Example: education levels; social cohesion	Measures must be specific for the determinant; e.g. literacy levels for education	Needs to be by region (as aggregation could lose information)	Linking these social determinants to health research is difficult and requires additional research	All pillars, but mainly pillar IV
	Environmental determinants	Example: air pollution levels	Level of known toxic pollutants in the air (parts per million)	Needs to be by region (as aggregation could lose information)	Dependent on environmental risk factor under study	All pillars, but mainly pillar IV
DETERMINANTS OF HEALTH	*Acceptability*	Example: self-reported patient satisfaction	Surveying patients to identify their experience of the health service	* Could be applied from health care provider level to regional * Not useful beyond regional levels since information would be lost in aggregation	Some self-report surveys are not rigorous data collection tools and should be used with caution	Particularly pillar III
	Accessibility	Example: wait times	Wait times for specific conditions and/or interventions	Useful at provider, region, or population levels	Only applicable to secondary care	Particularly pillar III
		Example: appointment statistics	Time to appointments for different groupings (e.g. socio-economic, gender, ethnicity)	Useful at provider, region, or population levels	Potentially difficult to access disaggregated statistics from physicians	Particularly pillar III

HEALTH IMPACTS						
Category	**Subcateg-ory**	**Indicator**	**Description**	**Level of Application**	**Comments**	**Pillars that indicators are relevant to**
	Appropriateness	Example: adherence to clinical guidelines	Identifying whether practice conforms to the most up-to-date evidence base	* Can be used in audit for individuals * For evaluation it is most useful at provider, region, or national levels	Requires an audit of clinical practice, which needs to be based on a standardized survey	Particularly pillar III
	Competence	Example: civil law suits against the health system	Counts of civil law suits by clinical area over time	Could be used to show data from individuals upwards, depending upon the defendant involved in the suit	Civil law suits only identify the most extreme examples of incompetence, but measures of competence itself are difficult to come by	Particularly pillar III
	Continuity	Self-reported continuity of care	Surveying patients to identify their perception of the continuity of their care	Could be applied for individuals, health care providers, or regions	Self-reported data relies on standardized data collection across Canada	Particularly pillar III
	Effectiveness	Example: re-admission rates	Numbers of re-admissions by condition over a set time period; year-on-year change	Useful at provider, region, or population levels	Can only provide information on conditions that require secondary care	Particularly pillar III
	Efficiency	Actual vs. expected hospital stay	Length of stay for a patient compared to the expected stay for the condition	Not useful for individuals, only for provider, region, or national comparisons	Only provides data on secondary care and cannot take into account individual complications or co-morbidity	Particularly pillar III
		Cost input versus output	* Data on the inputs to health care services and on the different factors identified as outputs (e.g. available beds, emergency admissions, etc.) * Can be fed into a stochastic model to identify efficiency	Provider-level analysis only	Much of the data for any analysis is already collected for health care providers	Particularly pillar III
	Safety	Example: adverse drug effects	Numbers of adverse drug effects; year-on-year change	Provider, provincial, and federal levels	Adverse drug effects are an easily measurable safety issue, and one of the most visible	Particularly pillar III
		Example: hospital-acquired infections	Levels of HAI; year-on-year change	Provider, provincial, and federal levels	HAIs are a very current safety issue and are easy to measure and link to specific policies and research findings	Particularly pillar III

BROAD ECONOMIC AND SOCIAL IMPACTS					
Category	**Indicator**	**Description**	**Level of Application**	**Comments**	**Pillars that indicators are relevant to**
ACTIVITY IMPACTS	Economic rent (Labour rents)	The economic benefit (in $) of employing people in health research rather than in another capacity	* May be applicable at the funder or disease area level * Most useful at a provincial/national level	More comprehensive than simple employment benefits since it accounts for the counterfactual of what individuals would do if they weren't involved in research	All pillars
COMMERCIALI ZATION	Licensing returns ($)	Dollars spent on licensing patents held by Canadian organizations/individuals	* Not recommended for individuals * Recommended for groups/ institutions/ provinces/ nationally	Can be linked to specific research findings	All pillars, likely to be emphasis on pillars I and II

BROAD ECONOMIC AND SOCIAL IMPACTS					
Category	**Indicator**	**Description**	**Level of Application**	**Comments**	**Pillars that indicators are relevant to**
	Product sales revenues ($)	Sales revenues of products developed in Canada	* Recommended for provinces and nationally – could be used for specific funders * Not recommended for individuals; groups or institutions	Difficulty in linking to research findings means not useful for assessing research groups	All pillars, likely to be emphasis on pillars I and II
	Valuation of spin-out companies ($)	Using the valuation of portfolios of new spin-out companies and the sales of spin-outs to provide the value to the economy of spin-outs at any given point (annually)	* Recommended for provinces and nationally – could be used for specific funders * Not recommended for individuals; groups or institutions	Accessing valuation of new spin-outs may be difficult but are presumably available through venture capital firms that support the spin-out companies	All pillars, likely to be emphasis on pillars I and II
	Economic rent (Producer rent and spillover effects)	* Producer rent is the economic benefit to a company on top of expected revenues * Spillover effects are the external effects of investing in R&D on groups not invested in (e.g. investment from abroad in private R&D having benefits in Canada)	* Recommended for provinces and nationally – could be used for specific funders * Not recommended for individuals; groups or institutions	Calculating producer rent and spillovers has been performed for health R&D, but requires understanding of economic techniques underpinning analysis	All pillars, likely to be emphasis on pillars I and II
HEALTH BENEFIT	Health benefit in QALYs per health care dollar	Improvement in health measured through QALYs gained and divided by the cost of achieving that health gain	* Not recommended for individuals or groups * Useful for institutions/funders/ provinces/ nationally	QALYs can be monetized (controversial methodology) so a monetary net benefit could be compared to other uses of capital	All pillars
	Health benefit in PROMs per health care dollar	Improvement in health measured through PROMs gained and divided by the cost of achieving that health gain	* Not recommended for individuals or groups * Useful for institutions/funders/ provinces/ nationally	PROMs have not been monetized so this measure can only be compared to other PROMs measures	All pillars
WELL-BEING	Annual report of HRSDC	Human Resources and Social Development Canada (HSRDC) has multiple indicators of well-being that can be used to identify well-being	* National level only as difficult to attribute changes to research findings	* No links to research (health or otherwise) except through the "health" section of the well-being indicators, which are covered in the Health Impacts category * Data already collected and publicly accessible	All pillars, emphasis likely on pillar IV
	Happiness	As measured using established survey techniques for happiness-depression	* Recommended for provinces and nationally * Not recommended for individuals, groups, or institutions	* Self-report happiness scales used by Statistics Canada * Very difficult to make any link to health research findings currently	All pillars
	Level of social isolation	Loneliness scales for measuring social isolation of individuals	* Recommended for provinces and nationally * Not recommended for individuals, groups, or institutions	* Tools exist for measuring * Very difficult to make any link to health research findings currently	All pillars
SOCIAL BENEFITS	Socio-economic status	Identifying socio-economic status of individuals in Canada	* Recommended for provinces and nationally * Not recommended for individuals, groups, or institutions	* Causality of socio-economic status to health outcomes is well known * Not understood if health research can alter socio-economic status * Collected to identify if changes in socio-economic status correlate with research impacts	All pillars, emphasis likely on pillar IV

`

One key fact that has emerged from this assessment is that the science behind defining returns on investment in health research is embryonic and thus presents a significant opportunity for advancement. When we began this assessment, there was considerable interest in the topic and some base to build on, but there was no established model or validated method for tracing the impacts of health research. We think that this assessment has cast some light on the topic and on the issues that need to be resolved in order to truly define "returns on investment from health research". The commissioned papers that are appended to this report also contain valuable insights into the topic, including considerations for the development of other potential indicators.

Based on the path forward suggested by this assessment, our panel hopes that the "science of health research" will be advanced, in part, by the use of the impact framework and menu of indicators and metrics developed here, and by adding to the indicators over time. The longer term impact of using the framework and indicators will hopefully be the successful identification of validated metrics of impact, resolution of at least some of the attribution problem, quantification of the returns on investment in health research with both economic and value-based societal outcomes, and more rapid improvement of health research impacts over time.

Based on our assessment, the panel made five recommendations.

Panel Recommendations

1. **The framework and indicators identified by this assessment should be used by all funders of health research in Canada for evaluation of their health research impacts.** The measurement framework developed during this assessment incorporates the positive attributes of current best practices internationally (reviewed in Chapter 2) and reflects the needs of all Canadian and international funders of health research. It can be used to address existing and future evaluation questions for the purposes of accountability, advocacy, and for learning. The framework also provides a summary overview of the main sources, paths, effects, and impacts of individual or cumulative strands of research. It identifies the stages in the progression from funding to knowledge production, dissemination, uptake, translational modification, adoption, and impacts on outcomes over time, all of which are essential for a full appreciation of the impacts of funded health research.

2. ***Sets* of indicators and metrics chosen from our menu should be used by all funders of health research in Canada for evaluation of their health research impacts.** As reviewed in Chapter 4, it is critical for impact evaluators to recognize that multiple indicators and metrics are required for any evaluation, because any single indicator or metric can be dismissed as partial, imperfect, and likely to distort findings. Different subsets of indicators and metrics may need to be selected by evaluators to address different evaluation questions, since no single subset of indicators is ideal for every impact evaluation.

   It should also be noted that the "science of health research impact indicators and metrics" is still embryonic. The impact indicators we have assembled provide only a starting menu of indicators as of 2008. There are several areas where indicators are required, but the underlying methodologies have yet to be developed. In other areas where the methodology is clear, the underlying data are not routinely collected. In general, as the indicators and metrics move from more proximal impacts (for example, relating to published research results) to more distal impacts (for example, impacts on population health and the economy), the underlying data and methodology grow weaker and require more development.

3. **Canadian health research funders should begin collaborations immediately to advance the practical (methodological) prerequisites for measuring returns on investment in Canada.** There are a number of specific areas where evaluation of the impacts of health research is hampered by varying practices across research funders. These include the variable nomenclature used to describe different fields and topics in health research, data on the health research process itself, and inconsistent or limited accounting for the resources actually consumed in undertaking health research. As described in chapter 4, there is also a need to create and expand a central resource of scientifically validated indicators (with evidence that is easily accessible to the world). These needs drive a number of sub-recommendations for this new collaborative:

   a. *A commitment is required from all stakeholders and funders to standardize and refine methods, and to routinely collect high quality and appropriate data*. Current gaps in capturing research outputs and impacts must be addressed at the outset of research being performed, with good data collection methods and sharing strategies (with incentives to provide and share good data). This includes making a commitment to:

collect standardized information on common CVs, perform consistent end-of-grant reporting, use common sampling methodologies, standardize case study methodologies, establish common definitions and classifications of research disciplines, and standardize key words, etc. Research communities and stakeholders should be involved in defining what questions are to be answered, what data are required to answer them, what incentives should be considered and what data collection methods should be endorsed to build databases. Their time commitments to data collection obviously should be minimized by developing the most efficient methods possible. They should then participate in collecting good data.

b. *A library of impact indicators and metrics should be created, beginning with the starting menu developed here and then adding to it (using the criteria identified*). This library should be maintained on an easily accessible website to provide an on-going resource for sharing definitions and interpretations of the best indicators and metrics and their use. It should be updated regularly, because the number and sophistication of indicators and metrics will continue to evolve as new methods, new denominators, or new sources for indicators are developed (for example, the web as a data source). This investment will help to avoid unnecessary duplication and wasted efforts from using invalid or non-comparable indicators within Canada. It will also provide a readily accessible platform from which to contribute internationally to making health research impact evaluation more comparable and effective.

c. *A core set of key health research impact questions – based on what is practical and feasible – should be developed in Canada using our framework and indicators.* Canadian health research funders, together with their communities and stakeholders, must identify impact evaluation questions and define what indicator sets are required to answer questions for their own purposes (such as accountability, advocacy, learning or some combination of these) in their respective organizations. They must also implement appropriate data collection processes to support the construction of required databases.

i. *Strategic and ethically sound selection of indicator sets is required in order to avoid biasing future health research.* Because the selection by funders of evaluation criteria generally, and indicator sets specifically, has significant steering effects on researchers, funders and governments, “preferred sets of indicators and metrics” must be chosen carefully—particularly those with long-term implications (such as outputs, adoption and outcomes).

ii. *Evaluation questions and the choice of indicators and metrics to be used reflect political and social choices; the motives underpinning such choices should be made transparent.* Stakeholders need to understand the rationale for choices and be able to debate them.

4. **Canada should immediately initiate a national collaborative effort to begin to measure the impacts of Canadian health research.** Rather than waiting for perfect data collection methods and ideal, comprehensive, and validated indicator and metrics sets to be developed (see Recommendation #3), there is a need to immediately begin to use the framework and indicators to measure the impacts of existing Canadian investments in health research. This requires that

leaders from national organizations, industry and government organize a comprehensive effort aimed at a combination of learning, accountability and advocacy evaluation objectives. Research communities, other stakeholders and members of the public must also be engaged in that exercise. One option is that funders, led by one national organization, form a national council to lead strategic planning and execution of the framework, with a formal secretariat and commissioned data collectors to begin this work.

5. **Canadian health research funders should collaborate internationally to advance the "basic science of health research impacts."** Given the many unresolved issues of attribution, time-lags, the counterfactual and data collection limitations discussed in Chapter 3, it is recommended that Canada establish international collaborations to advance the "basic research" in this field. Such collaborations with international experts will help to:

    a. *resolve complexities:* (i) the diverse fields of health research generate a wide range of outputs that affect health, wealth, and well-being; (ii) advances in health are rarely (if ever) attributable to a specific funder or a specific discovery; (iii) the components of health research interact with one another; (iv) the uptake of health research is influenced by many complex factors including the incentives placed on potential adopters of research, the political environment, perceived and real inequities, and many other factors within the context into which research results are added.

    b. *improve the global system of health research impacts:* Understanding the world system of knowledge translation can help to improve it, by identifying key steps, eliminating bottlenecks, improving models to integrate across pillars, and testing other knowledge system innovations as they are discovered.

    Recommended plans include:

    i. *an international funding stream* – Collaborative international funding would be beneficial in advancing the elements identified in Recommendation #3.

    ii. *a research plan to identify "contribution indicators" and close "attribution gaps"* – The distal indicators of the impacts of health research are very broad, and many factors other than funded health research can and do affect these outcomes. It is fundamental that methodologies be developed to separate the contribution of health research from other causal factors.

    While it will never be possible to quantify all health research impacts with precision and accuracy, it is feasible that gaps can be closed by a concerted research effort to do so. (Caution: While gaps are being closed, all funders and practitioners must be aware of the issue of attribution.)

    iii. *a research plan to use the recommended framework for learning (impact improvement) purposes* – As noted above, the framework can be used to study where, when, how, and why knowledge may or may not be translated over time. A systematic approach to evaluating knowledge flows, barriers and facilitating factors that influence outputs and outcomes can be defined over time. Learning will help to guide better health-research investment decisions in the future.

Assessment Panel Members

- Cyril Frank, MD (Chair), McCaig Professor of Joint Injury and Arthritis Research; Executive Director of the Alberta Bone and Joint Health Institute; and Professor, Division of Orthopaedics University of Calgary, Calgary, AB, Canada
- Renaldo Battista, MD, MPH, ScD, Professor and Director of the Department of Health Administration, Université de Montréal, Montreal, QC, Canada
- Linda Butler, Fellow and Head, Research Evaluation and Policy Project, Australian National University, Canberra, Australia
- Martin Buxton, BA, Professor, Health Economics, Brunel University, United Kingdom
- Neena Chappell, PhD, Canada Research Chair; Social Gerontology, Professor of Sociology and Centre on Aging, University of Victoria, Victoria, BC, Canada
- Sally C. Davies, Director General, Research and Development, Department of Health and National Health Service, United Kingdom
- Aled Edwards, PhD, Banbury Professor, Banting and Best Department of Medical Research, University of Toronto, Toronto, ON, Canada
- Chris Henshall, PhD, Pro-Vice-Chancellor, External Relations, University of York, York, England, United Kingdom
- Yann Joly, LLB, LLM, project manager, Centre de recherche en droit public Université de Montréal, Montreal, QC, Canada
- Gretchen Jordan, PhD, Principal Member of Technical Staff, Science and Technology Strategic Management Unit, Sandia National Laboratories, Department of Energy, Washington, DC, United States
- Terence Kealey, MB, BS, PhD, Vice Chancellor & Clinical Bio-Chemist, University of Buckingham, Buckingham, United Kingdom; author *The Economic Laws of Scientific Research*
- Michael C. Wolfson, PhD, Assistant Chief Statistician, Analysis and Development, Statistics Canada, Ottawa, ON, Canada
- Steven H. Woolf, MD, MPH, Professor, Departments of Family Medicine, Epidemiology, and Community Health, Virginia Commonwealth University, Richmond, VA, United States

Sponsors of The Report

Major Sponsors

- Canadian Health Services Research Foundation (CHSRF)
- Canadian Institutes of Health Research (CIHR)
- Canada's Research-Based Pharmaceutical Companies (Rx&D)
- Public Health Agency of Canada (PHAC)

Sponsors

- Alberta Heritage Foundation for Medical Research (AHFMR)
- Association of Canadian Academic Healthcare Organizations (ACAHO)
- Association of Faculties of Medicine of Canada (AFMC)
- BIOTECanada
- Canadian Agency for Drugs and Technologies in Health (CADTH)
- Fonds de la recherche en santé du Québec (FRSQ)
- Government of Ontario, Ministry of Research and Innovation; Ministry of Health and Long-Term Care
- Heart & Stroke Foundation of Canada (HSFC)
- Manitoba Health Research Council (MHRC)
- Michael Smith Foundation for Health Research (MSFHR)
- National Cancer Institute of Canada (NCIC)
- Nova Scotia Health Research Foundation (NSHRF)
- Ontario Neurotrauma Foundation (ONF)
- Saskatchewan Health Research Foundation (SHRF)
- Western Economic Diversification Canada (WD)

Contributors

- Canada Foundation for Innovation (CFI)
- Canadian Association of Schools of Nursing (CASN)
- Canadian Medical Association (CMA)
- Canadian Nurses Association (CNA)
- Canadian Nurses Foundation (CNF)
- Newfoundland & Labrador Centre for Applied Health Research (NLCAHR)
- Research Canada