BowelScreen
Data Quality Statement

Data Quality Statement to accompany data released by BowelScreen. It includes data quality assurance, contextual information, the methods used to compile the statistics, and other background information that data users may find useful.

An tSeirbhís Náisiúnta Scagthástála
National Screening Service

BowelScreen
The National Bowel Screening Programme

BSP/DQ-4 Rev 01
Published 2025

Table of Contents

1 Introduction

BowelScreen is a population-based bowel screening programme. It offers free bowel screening to men and women within the eligible age range (currently age 58 to 70) who are on our register. BowelScreen puts together a register of men and women eligible for screening (see section 5.1). People who are on our register are offered screening every two years while they are within the eligible age range. We aim to offer people their first invitation between the ages of 58 and 60. The age range for bowel screening is expanding. Over time bowel screening will be offered to people aged 55 to 74[1].

Bowel Screening is performed via a self-administered home test known as a faecal immunochemical test (FIT). The FIT is based on the detection of a level of blood in the faeces (stool)[2].

BowelScreen maintains a register of people within the screening age range. This information is held by BowelScreen as a *population register*. BowelScreen also collects selected data relating to BowelScreen participants (Appendix A). The population register data and other selected data are securely stored on the BowelScreen database (known colloquially as 'COR').

2 Purpose

A Data Quality Statement (DQS) is a feature of HIQA's national standards for information management[i]. A DQS is a statement prepared to accompany data releases from national data collections that highlight the dimensions of data quality, including strengths and weaknesses[3]. The purpose of this DQS is to accompany data release(s) from the National Screening Service (NSS) regarding BowelScreen. This document provides further information and/or context regarding the accompanying data release. This document details how BowelScreen uses data and specifically:

- Why the NSS can legally collect this data
- Where the NSS stores the data
- How the NSS collects data and from whom it is collected
- Who the data in the data release refers to (the dataset cohort)
- How the NSS calculates any figures within the data release
- How the NSS defines and assures data quality
- What caveats the NSS has applied to the data release for potential amendments

3 Legislative remit for data collection

BowelScreen has a legislative remit to acquire its population register information under the Health (Provision of Information) Act, 1997 "*for the purposes of compiling and maintaining a record of the names, addresses and dates of birth of persons who, for public health reasons, may be invited to participate in that programme*"[4][5].

This register data is *personal data*[6]. In the provision of the national bowel screening programme, BowelScreen processes this *personal data*. The law on protecting personal information, known as GDPR, allows BowelScreen to use personal information to invite people for screening[7]. As part of the NSS (and part of the Health Service Executive), BowelScreen carries out this processing of *personal data* on the lawful basis that the processing is necessary for a task carried out in the public interest, or in the exercise of the official authority officially given (vested in) the controller. For the HSE this official authority is vested in the NSS through the Health Act 2004[8].

Upon accepting/uptake of a screening invitation, BowelScreen collects and processes certain *special category data*[9] relating to service users, specifically health data required to deliver the programme. The law on protecting personal information, known as GDPR, allows BowelScreen to process health

[i] Standard 3.2.2: *Data quality reports* which detail findings of *data quality assessments*, including *data quality statements.*

data[10]. BowelScreen is a body of the NSS (and part of the HSE). As such, BowelScreen carries out this processing of *special category data* on the basis that:

- It is necessary in the provision of healthcare or treatment.
- It is necessary for medical diagnosis.
- It is utilised for the management of BowelScreen by the NSS for the provision of a quality assured screening service.
- It is processed by either a healthcare practitioner or a person who, in the circumstances, owes a duty of confidentiality to the data subject that is equivalent to that which would exist if that person were a health practitioner.
- The processing is necessary for a task carried out in the public interest or in the exercise of official authority vested in the controller.

The law on protecting personal information, known as 'GDPR', allows BowelScreen to use personal information to ensure that a safe and effective screening programme is provided across the Republic of Ireland[7][10][11].

4 BowelScreen database (COR)

BowelScreen data is stored in a SQL Server database known colloquially as 'COR'. COR is a bespoke, event-driven application which respects a logical order and follows the development of the screening process. All data collected and maintained on COR has a specific purpose in the provision of a quality assured screening service. BowelScreen collects an agreed dataset for all parts of the screening pathway. This includes consent; invitation; the screening test; diagnosis; surveillance; and cancer treatment. The data collected is linked to the BowelScreen charter commitments and the BowelScreen standards for quality assurance in colorectal screening[12], which are based on the EU Guidelines for colorectal cancer screening[13].

5 Data sources

The statistics in the accompanying data release are derived from information that is routinely collected by the NSS for the operation of BowelScreen, including for quality assurance and performance management purposes.

5.1 Population register

Information on the COR database regarding the population register is supplied from:

- The Department of Social Protection (DSP)
- Online self-registration or self-registration by freephone[ii]

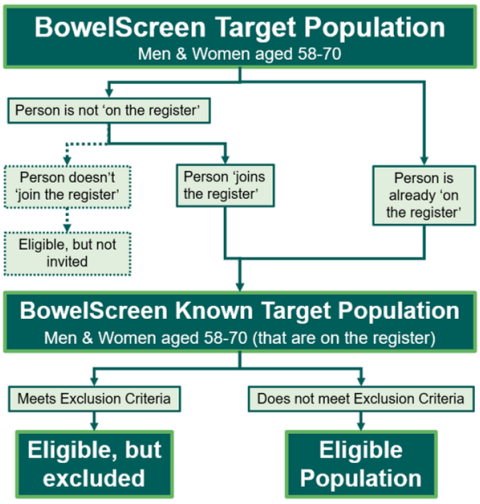
Data supplied to the NSS for the BowelScreen population register is subject to data processing for the purpose of data cleansing/standardisation and duplicate identification. This involves both a file import process from the Department of Social Protection (DSP) and self-registration.

5.2 COR database

The client data, FIT data, appointment data, and any subsequent clinical data (such as endoscopy, surgery, etc.) is collated within the BowelScreen COR database. The data on COR is routinely collected as a participant progresses through various pre-defined pathways within BowelScreen. Progress through various pathways within the COR information system is driven via 'events'.

[ii] Freephone 1800 45 45 55

6 Data cohort & scope



As of October 2025, the BowelScreen target population is people aged 58 to 70 years old, living in Ireland. People aged 58 to 70 years old living in Ireland need to be on the BowelScreen register in order to be part of the *known target population*[iii], and thus sent an invitation to take a screening test. People aged 58 to 70 years old living in Ireland fall outside the *eligible population*[iv] if they are on the BowelScreen register but are '*excluded*'. Screening eligibility criteria are documented internally[v] within BowelScreen.

6.1 Cohort selection for the data release

Cohort selection for the BowelScreen data release is based on those who were invited for bowel screening within the calendar year or screening round (depending on the reporting period). Therefore, the figures in BowelScreen data releases relate to people invited by BowelScreen within the relevant calendar year. Some of these people may have been screened or treated in the following year.

6.2 Cohort selection for PCCRC rate

Cohort selection for the calculation of the Post-Colonoscopy Colorectal Cancer (PCCRC) rate differs from the cohort selection outlined in section 6.1. The cohort selection criteria for the PCCRC rate is detailed within the data release.

7 Methods used to compile the statistics.

The NSS analyses the data using reports developed in Microsoft SQL Server Reporting Services (SSRS). The statistics in the accompanying data release are compiled by the NSS Programme Evaluation Unit (PEU).

The figures are published in downloadable PDF format and are available online. The figures presented in the data release are in the form of simple counts, percentages (rounded to one decimal place) or rates (e.g., number of colorectal cancers detected per 1,000 screened). Definitions and formulae detailing how the statistics in the data release are calculated can be found in the BowelScreen standards for quality assurance in colorectal screening (3rd ed.).

[iii] The *known target population* is all people of screening age living in Ireland that are known to the programme.
[iv] The *eligible population* is the *known target population* minus those people that are *excluded*.
[v] Q-Pulse Document Reference: BSP/QP-001 *'BowelScreen Eligibility Exclusions'*

8 Data quality

Quality data is data which is *'fit for purpose'*. The quality of data (fitness for purpose) is determined through assessment against internationally accepted dimensions. Data quality dimensions need to be balanced against one another to meet users' needs. These dimensions are relevance; accuracy & reliability; timeliness & punctuality; coherence & comparability; accessibility & clarity[14][15][16].

8.1 Relevance

Relevant data meets the current and potential future needs of the data users. Data is deemed relevant when it is **useful** and **valuable**[17].

The information in the accompanying data release is relevant since it is **useful** for demonstrating compliance with key performance indicators, provides accountability to the taxpayer and demonstrates the **value** of the programme.

According to HIQA, **data users are stakeholders** that use data held by a screening programme or use data outputs from that programme[3]. BowelScreen has clearly defined *data users* and *data uses*. The data users for the statistics contained within the data release include those outlined in the NSS stakeholder engagement framework. The relevance/utility/use of BowelScreen data is documented and categorised here.

8.2 Accuracy & reliability

The **'accuracy'** of data refers to how closely the data correctly describes what it was designed to measure. Reliability refers to whether that data consistently measures, over time, the reality that it was designed to represent. The accuracy & reliability of the data in the accompanying data release is determined through its data **coverage, collection, completeness, processing,** and **revisions** [17].

Data **'coverage'** refers to the degree to which the data available covers the target population[17]. The PEU measures *target population* coverage by benchmarking a snapshot of the BowelScreen population register against census data from the Central Statistics Office (CSO) on the date of a census. Duplicate record identification is a part of the client upload process. The *eligible population* that has been invited and screened is measured in the quality assurance standards 3.1 and 3.2 respectively.

'**Collection'** refers to the procedures that are in place to ensure that the data is captured in a usable format[17]. Data for BowelScreen is collected through a combination of file uploads and manual data entry. A guide to data entry and use of COR is provided to BowelScreen data users for the clinical portal[vi], the endoscopy reporting system (ERS)[vii] and the population register[viii]. BowelScreen data is collected in structured formats (e.g. boolean, integer, decimal, date/time) with unstructured string data kept to a minimum and only used where necessary. The format of BowelScreen variables is documented in the BowelScreen data dictionary.

'**Completeness'** refers to the degree to which individual variables are present within a dataset[17]. BowelScreen provides a self-registration service to people online and by freephone as part of efforts to attain 'completeness' regarding register data. PEU measures and monitors the completeness of the BowelScreen population register through comparison with national census data from the Central Statistics Office (CSO). Completeness of clinical data is achieved through a combination of mandatory fields, user prompts, and validation messages. This is documented in the *'verification rules'* section of BowelScreen's data dictionary.

[vi] Q-Pulse Document Reference: BSP/END-005 *'COR Data Entry Clinical Portal'*
[vii] Q-Pulse Document Reference: BSP/END-006 *'Endoscopy Reporting System (ERS) User Manual'*
[viii] Q-Pulse Document Reference: BSP/OPS-005 *'Call Centre Daily Operation of BowelScreen Population Register'*

'**Processing'** is the transformation of data from the form in which it is received into another form that facilitates analysis[17]. Data processing is carried out on the population register to identify duplicates. Data processing is also carried out on COR data to produce *clinical reports* and *programme evaluation reports*. This data processing is carried out using Microsoft SQL Server Reporting Services (SSRS).

8.3 Timeliness & punctuality

Good quality data should be both **timely** and **punctual**[17].

Timely data is collected within a reasonable agreed time period after the activity that it measures[17]. BowelScreen data releases from the NSS are published some months after the end of a calendar year. This is because:

- Cohort selection within the data release is based on client *'invited'* date. As such, it can take some time after the end of the calendar year to gather a complete dataset for the cohort (particularly for those invited in the final quarter of the calendar year and those who have been referred for treatment)
- Trade-offs exist between data quality dimensions (timeliness vs. accuracy)[3].

Punctual data is data delivered or reported on the dates promised, advertised or announced[17]. As part of adherence to HIQA information management standards[ix], the NSS is committed to developing a formal data release calendar for all its screening programmes.

8.4 Coherence & comparability

Coherent and comparable data is consistent over time and across providers and can be easily combined with other sources. This data quality dimension is measured across **standardisation, coherence, historical comparability, and regional (Countries) comparability**[17].

'Standardisation' refers to the degree to which data is collected using common definitions or standards. Standardisation is assisted through use of *technical standards*, *messaging standards*, *terminology/classification standards*, and *metadata*[17].

Technical Standards: BowelScreen data is collected in standardised formats (e.g., boolean, integer, decimal, date/time) with minimum freetext data. Date/time data is stored in ISO 8601 standard format but that data is entered, extracted, and reported in the European IS/EN 28601 standard format.

Messaging Standards: BowelScreen does not currently utilise *Messaging Standards* but is working towards implementing the messaging standard Health Level 7 (HL7).

Terminology/Classification Standards: BowelScreen does not utilise *Terminology & Classification Standards* within its database. However, as part of adherence to HIQA information management standards[x], each of the variables in the BowelScreen database has been implemented according to various national/international recommendations & guidelines. This standardisation is documented internally[xi]. Examples include; Paris Classification, Modified Gloucester Scale, TNM Staging, RCPath dataset (G049), AJCC Staging Standard, and C-RADS.

Metadata: As part of adherence to HIQA information management standards[xii], BowelScreen has documented metadata in the form of a data dictionary. The structure of BowelScreen's data dictionary is based on HIQA guidance, and can be viewed in Appendix B. The HSE has established the Dataset Specification Management Process (DSMP) to ensure a consistent approach to dataset standardisation. The HSE is also developing a data dictionary based on the ISO/IEC 11179 Metadata Registry (MDR) standard.

[ix] Standard 2.1.1: Providing access to high-quality information to promote effective and safe use of data in a *timely* manner.
[x] Standard 3.1.2: Compliance with *health information standards* (e.g., messaging standards, terminology standards, IHI,)
[xi] Q-Pulse Document Reference: BSP/DQ-1 Rev01
[xii] Standard 3.1.3: Developing and implementing a *data dictionary* to ensure consistency and comparability.

'Coherence' refers to the degree to which it is possible to combine and make use of related data from different sources[17]. As previously outlined, BowelScreen data is collected in a manner coherent with those outlined in European and National standards. BowelScreen data is combined with NCRI data to determine the Post-Colonoscopy Colorectal Cancer (PCCRC) rate. As part of adherence to HIQA information management standards[x], there are plans for the implementation and use of Individual Health Identifiers (IHI). The use of IHI within the NSS is legislated for in the 2014 Health Identifiers Act[18]. The implementation of IHI will improve the efficiency of data linkage with other datasets in Ireland, leading to a more integrated and coherent approach to the use of health information[19][20][21][22].

'Historical comparability' refers to how the consistent use of definitions and standards facilitates comparison over time[17]. Historical changes to BowelScreen's data (such as inactive variables, deactivated values, etc.) are logged in the BowelScreen data dictionary. There has been a historical change to BowelScreen's FIT cut-off score. There has also been historical changes to BowelScreen's eligible age range. Data users should be aware of these historical changes (Appendix C).

'Regional comparability' refers to how the consistent use of definitions and standards facilitates comparison across countries[17]. Data users should be aware that bowel screening programmes in different countries can employ different *'cut-off'* scores (sometimes referred to as *'threshold score'*) for FIT results. These *'cut-off'* scores can also be in different units of measurement across countries. The *'cut-off'* score is important because the higher the cut-off score, the lower the positivity rate and the higher the specificity and positive predictive value (PPV) for colorectal cancer (CRC) and advanced adenomas[23].

8.5 Accessibility & clarity

Good quality data is easily obtainable and clearly presented in a way that can be understood easily; it is **accessible** and **interpretable**[17]

'Accessibility' refers to the ease with which data can be identified, obtained, and used[17]. The accompanying data release is digitally accessible in PDF format. As part of HIQA information management standards[xiii], BowelScreen has made a publications catalogue available online.

'Interpretability' refers to the degree to which users are provided with the required documentation and metadata to assist them in understanding the data[17]. As part of adherence to HIQA information management standards[xiv], BowelScreen metadata is documented in its data dictionary. BowelScreen Quality Assurance Standards contains metadata regarding counts and calculations. This document also contains metadata regarding historical changes.

[xiii] Standard 4.2.1: Optimising the accessibility, use and value of information through effective sharing and dissemination.
[xiv] Standard 3.1.3: Developing and implementing a data dictionary to ensure consistency and comparability.

9 Confidentiality, transparency & security

9.1 Confidentiality

In order to maintain confidentiality and protect information from misuse, all members of staff within the NSS have five main responsibilities:

- Storing & transferring information securely
- Restricting access to a data subject's personal data[xv]
- Maintaining internet and email security[xvi]
- Improving data collection, data quality, and reporting
- Reporting data breaches[xvii]

The NSS works in collaboration with the NCRI to identify post-colonoscopy colorectal cancers (PCCRC) for BowelScreen. This work involves data sharing/processing.

As part of adherence to HIQA information management standards[xviii], a *data protection impact assessment* (DPIA) has been carried out on the process of processing/sharing data with the NCRI. The DPIA outlines the legal basis for processing the data, the legal basis for processing *special category data*, role of the data controller, any automated decision making, recipients of the *personal data*, identified risks & threats, technical security measures and controls in place.

9.2 Transparency

As part of adherence to HIQA information management standards[xix], a *privacy statement* is in development within the NSS. A link to the *HSE privacy notice* is provided here. The sharing or exchange of personal data between the NSS and the NCRI is governed by a data sharing agreement (DSA).

As part of adherence to HIQA information management standards[xx], BowelScreen has a *statement of purpose* which is publicly available here.

9.3 Security

The NSS utilises specialised firewall software and hardware to protect the NSS network and the data and equipment within it from security risks. NSS laptops have encryption software that ensures the data-at-rest is protected from theft or loss. Various security applications are used to protect NSS systems and data in a number of ways including providing anti-virus protection, uncovering hidden malware, and scanning web traffic for infected websites that can install malware or be used as gateways for cyber-attacks. The sharing or exchange of data between the NSS and the NCRI is carried out securely through a cross-platform interface engine.

BowelScreen programme data (COR) is stored in a SQL Server database. '*Live*' and '*Report*' versions of the database are kept on virtual servers. In addition to this, a backup of the live database is maintained on a separate virtual server. The servers are within the NSS network and are protected from security vulnerabilities. Access to the data within the network is via a physical network connection or via a virtual private network (VPN) which provides enhanced security. Access to the BowelScreen database is provided on an as-needed basis via username and passwords which are issued based on manager approval.

[xv] *Fundamentals of GDPR* training and certification is mandatory for NSS staff.
[xvi] *Cyber Security Awareness* training and certification is mandatory for NSS staff.
[xvii] Q-Pulse Document Reference: NSS/SOP-GDPR1 *'GDPR Data Breach Notification Escalation Process'*
[xviii] Standard 1.2.3: Undertaking *Data Protection Impact Assessments* to identify and mitigate any data protection-related risks.
[xix] Standard 1.2.2: Publishing and regularly reviewing a *privacy statement* or *privacy notice*.
[xx] Standard 4.1.6: Publishing a *statement of purpose* in an accessible format.

10 Data revisions

Statistics published in the accompanying data release are final statistics.

11 Strengths & weaknesses

Strengths	Weaknesses
Relevance: The users and utility of the data is documented. The data presented in the data release is deemed relevant to data users since it measures quality parameters.	**Punctuality:** Punctuality is not currently measured due to lack of *data release calendar*. The NSS is working towards a data release calendar for all its programmes.
Transparency: A *statement of purpose* is freely available online for data users.	**Transparency:** A *privacy statement* is not currently available. The NSS is working towards a *privacy statement*. The HSE privacy notice is here
Standardisation (standards): National & EU guidelines determine the standardisation of BowelScreen data. This is documented internally.	**Standardisation (standards):** BowelScreen does not utilise *Terminology and Classification Standards* (e.g., SNOMED, LOINC or ICD-10) to ensure semantic interoperability.
Accessibility: BowelScreen data releases are freely accessible online and a publication catalogue is available to data users.	**Accessibility:** The accompanying data is digitally accessible in PDF. This means that web-based tools are not applicable for manipulation.
Interpretability: In addition to its internal data dictionary, BowelScreen provides further metadata (numerator, denominator, etc.) in the Quality Assurance Standards.	**Interpretability:** BowelScreen's data dictionary is not currently publicly available.
Comparability: Historical changes to BowelScreen data are documented in its data dictionary. Historical changes to FIT threshold are documented in this statement.	**Comparability:** There is variance for regional comparability with other bowel screening programmes due to differences in age range and FIT threshold scores. This is an international issue.
Standardisation (metadata): BowelScreen utilises a data dictionary, the structure of which is outlined here.	
Coverage & completeness: The population register is measured routinely against census data by the Programme Evaluation Unit.	
Coherence: There are plans to utilise IHI in the BowelScreen database to facilitate an integrated and coherent approach to the use of health data.	
Accuracy & reliability: Automated and manual data processing procedures applied to the population register to minimise duplicate records.	

References

1 *BowelScreen Screening Charter*. The National Screening Service, BowelScreen
BowelScreen Charter (bowelscreen.ie)

2 *Health Technology Assessment (HTA) of a Colorectal Cancer Screening Programme in Ireland*, Health Information and Quality Authority, 2009
HTA of a population based colorectal cancer screening programme in Ireland (hiqa.ie)

3 *Guidance for a Data Quality Framework for Health and Social Care*, Health Information and Quality Authority, 2018
Guidance on a Data Quality Framework for Health & Social Care (hiqa.ie)

4 *Recommendation for Future Development of Colorectal Cancer Surgery in Ireland and Guidelines for the Management of Rectal Cancer in Ireland,* Irish Association of Coloproctology, Dublin; 2010
Future Development of CRC Surgery & Rectal Cancer Guidelines in Ireland (rcsi.com)

5 *Guidelines for the Implementation of a National Quality Assurance Programme in Gastrointestinal Endoscopy, version 2.0,* RCPI and RCSI, Dublin; 2011
Guidelines for Implementation of a National QA Programme in GI Endoscopy (rcpi.com)

6 ***GDPR Article 4(1)*** *'personal data means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person'*
Art. 4 GDPR – Definitions (gdpr-info.eu)

7 **GDPR Article 6(1)(e)** *'processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller'.*
Art. 6 GDPR – Lawfulness of processing (gdpr-info.eu)

8 *Privacy Notice, section 3,* Health Service Executive, 2020
HSE Privacy Notice (hse.ie)

9 ***GDPR Article 9(1)*** Special category data is '*personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation'.*
Art. 9 GDPR – Processing of special categories of personal data (gdpr-info.eu)

10 **GDPR Article 9(2)(h)** '*processing is necessary for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services on the basis of Union or Member State law or pursuant to contract with a health professional'*
Art. 9 GDPR – Processing of special categories of personal data (gdpr-info.eu)

11 **GDPR Article 9(2)(i)** *'Processing is necessary for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices, on the basis of Union or Member State law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy'*
Art. 9 GDPR – Processing of special categories of personal data (gdpr-info.eu)

12 *Guidelines for Quality Assurance in Colorectal Screening, Third Edition,* BowelScreen, Dublin; 2024
BowelScreen Guidelines for Quality Assurance in Colorectal Screening (screeningservice.ie)

13 *European Guidelines for Quality Assurance in Colorectal Cancer Screening and Diagnosis, First Edition,* European Commission, Belgium; 2011
EU Guidelines for Quality Assurance in CRC Screening & Diagnosis (bowelscreen.ie)

14 *Information Management Standards for National Health and Social Care Data Collections*, Health and Information Quality Authority, 2017
Information Management Standards (6 themes, 10 Standards) for data collections (hiqa.ie)

15 *Guidance on Information Governance for health and social care services in Ireland,* Health and Information Quality Authority, 2012
Guidance on Information Governance for Health Services in Ireland (hiqa.ie)

16 *National Standards for Information Management in Health & Social Care*, Health Information and Quality Authority, 2024
National Standards for Information Management in Health & Social Care (hiqa.ie)

17 *Data Quality Assessment Tool for health and social care,* Health Information & Quality Authority, 2018
Data Quality Assessment Tool for health and social care (hiqa.ie)

18 *Health Identifiers Act 2014 (No. 15 of 2014). Section 11: Use of individual health identifier and National Register of Individual Health Identifiers.* 8th July 2014. Act of the Oireachtas. Irish Statute Book
Health Identifiers Act 2014, Section 11 (irishstatutebook.ie)

19 *Evidence Synthesis: Recommendations on a consent model for the collection, use and sharing of health information in Ireland*, Health Information & Quality Authority, 2021
Recommendations on consent for use and sharing health information (hiqa.ie)

20 *Health Information in Ireland: A Socio-technical Analysis*, Health Research Board, 2015
Health Information in Ireland (ucd.ie)

21 *International review of approaches countries have taken to integrate National Health and Social Care Data Collections*, Health Information and Quality Authority, 2014
Recommendations for Integration of National Data Collections (hiqa.ie)

22 *Frequently asked questions about Health Identifiers*, Health Information and Quality Authority, 2015
FAQ on Health Identifiers (hiqa.ie)

23 *Expert Reference Group, Interval Cancer Report,* National Screening Service, Dublin; 2020.
BowelScreen Expert Reference Group Interval Cancer Report 2020 (hse.ie)

24 *Overview of Health Interoperability Standards*, Health Information and Quality Authority, 2013
Overview of Healthcare Interoperability Standards (hiqa.ie)

25 O'Donoghue D, Sheahan K, MacMathuna P, Stephens RB, Fenlon H, Morrin M, Mooney J, Fahy L, Mooney T, Smith A. *A National Bowel Cancer Screening Programme using FIT: Achievements and Challenges*. Cancer Prevention Research, 12(2), 89-94, 2019
National Bowel Cancer Screening Programme using FIT: Achievements & Challenges (accrjournals)

26 *Catalogue of National Data Collections*, Health Information and Quality Authority, 2022
Catalogue of National Data Collections (hiqa.ie)

27 *Screening Programmes: A Short Guide*, World Health Organisation, 2020
A Short Guide to Cancer Screening (hse.ie)

28 *General Data Protection Regulations: Frequently Asked Questions,* Health Service Executive, 2020
FAQ on General Data Protection Regulations (hse.ie)

Appendices

Appendix A – Data utility

Data Set	Data sub-set	Utility	Charter/ KPI / QA / Guideline
Person	Personal details Contact details Address details GP details Registration details Client status Client events	Record management, Communication, Decision making, Tracking eligibility, Pathway management, Charter commitments, KPI metrics	See section 3 of the BowelScreen Quality Assurance Standards.
Faecal Immunochemical Test (FIT)	Provider/supplier Barcode 'Lot' details Lab details FIT score & cut-off FIT kit events	Record management, Communication, Decision making, Pathway management, Charter commitments, KPI metrics	See section 4 of the BowelScreen Quality Assurance Standards.
Invitation	Invitation Invitation event	Record management, Communication, Pathway management, Charter commitments, KPI metrics	See section 3 of the BowelScreen Quality Assurance Standards.
Pre-assessment	Clinic / Location Round Surveillance Pre-Assess. event	Record management, Communication, Decision making, Pathway management, Charter commitments, KPI metrics	See section 5 of the BowelScreen Quality Assurance Standards.
Appointment	Appointment date Clinic / Location Prep type Endoscopy type Appointment event	Record management, Communication, Scheduling, Pathway management, Charter commitments, KPI metrics	See section 5 of the BowelScreen Quality Assurance Standards.
Endoscopy	Clinic / Location Date details Bowel prep details Endoscopy details Polyp details Tumour details Tattooing Sedation details Management plan Endoscopy events	Clinical details, Record management, Communication, Decision making, Pathway management, KPI metrics	See section 5 of the BowelScreen Quality Assurance Standards.
Post-colonoscopy results	Identifiers Lab Pathologist(s) Date details Polyp details Dysplasia Invasion Resection Staging Results event	Record management, Decision making, Pathway management, KPI metrics, Disease extent, Disease classification	See section 7 of the BowelScreen Quality Assurance Standards.

Data Set	Data sub-set	Utility	Charter/ KPI / QA / Guideline
CT colonography	Offered/CTC date Clinic Radiation dose CTC indication Completion Report date Radiologist(s) Adverse event(s) Report summary No. of findings CTC event	Record Management, Scheduling, Decision making, Pathway management, KPI metrics, Disease extent, Disease classification	See section 6 of the BowelScreen Quality Assurance Standards.
CTC findings	Finding type Finding size Finding morph. Finding location		
Neoadjuvant therapy	Location Outcome Radiotherapy Chemotherapy Dates Therapy event	Record management, Scheduling, Decision making, Pathway management, Charter commitments, KPI metrics	See section 8 of the BowelScreen Quality Assurance Standards.
Surgery	Date details Clinic / Location Surgeon Indication Surgery approach Surgery type Post-surgery leak Surgery event	Record management, Decision making, Pathway management, Charter commitments, KPI metrics, Disease extent, Disease classification	See section 8 of the BowelScreen Quality Assurance Standards.
Post-Surgery Results	Identifiers Lab/Hospital Pathologist(s) Date details Lymph node detail Site detail Adenocarcinoma Differentiation Margin Venous invasion Neoadjuvant Resection Staging Result event	Record management, Decision making, Pathway management, KPI metrics, Disease extent, Disease classification	See section 7 of the BowelScreen Quality Assurance Standards.
Meeting	Meeting date Clinic/Location Meeting type Decision Risk level Meeting event	Record management, Decision making, Tracking eligibility, Pathway management, KPI metrics	See section 8 of the BowelScreen Quality Assurance Standards.

Appendix B – Data dictionary structure

BowelScreen data dictionary structure	
Variable ID number	A unique reference number assigned to the variable. Each variable has a unique number assigned to it. This allows for specificity when discussing a variable.
Column ID number	A reference number for the column assigned to a variable within a table. It is an ordered number assigned to the variable as it appears on its table from left to right. This is sometimes referred to as the *"ordinal position"*.
Actively collected	Indicates if a variable is actively collected by BowelScreen. Where a variable has ceased to be collected, the date it ceased to be collected is indicated in the *"comments & guidelines"* section of the data dictionary.
Variable name [1]	The column header name of the variable within the database. This the name of the variable in the back end of the database. This name may contain no spaces.
Variable name [2]	The colloquial name for the variable within the database. This usually matches with the name given on the display screen (the font-end). This name may contain spaces.
Codes & values	Indicates the codes and values that are captured as part of the variable. This indicates whether a variable is Boolean, date, integer, decimal or string. Where the value is numeric with a link to a *"look-up table"* for the meaning of that value, then this section of the data dictionary will specify both the value and its associated meaning.
Look-up table (where applicable)	The name of the *"look-up table"* which is used to view numeric data with its associated meaning. This does not apply to all variables (e.g., date or Boolean variables will not have an associated *"look-up table"*.
Data type & field size	Refers to the type of data that is recorded for each variable. For BowelScreen this is TIMESTAMP, VARCHAR, NVARCHAR, BIT, INTEGER, DATETIME or DECIMAL. The data type column of BowelScreen's data dictionary also includes the field size for each variable. For example: DECIMAL*(19,5)* or *VARCHAR(40)*
Table	The table on which the variable is located. This usually indicates the fundamental dataset, but various tables can be combined to form datasets.
Definition	A definition which clearly explains what is meant by the variable
Related variables	Any other variables that are linked to or closely related to the variable. This section does not contain details of primary, foreign, or composite keys.
Comments & guidelines	This section contains relevant comments or guidelines related to the variable. It may contain a broad range of information (e.g., links to guidance documents, dates new values were introduced, further detail about the variable, changes to the variable over time, etc.)
Context	This section outlines the context in which the variable is collected
Verification rules	This section outlines the ways in which the accuracy of the variable can be verified. Where language such as *"should"* is used, this indicates that the system does not enforce the rule. Definitive terms like *"will"* indicate enforcement within the system.

Appendix C – Historical changes

FIT score
The FIT score indicates the score attributed to a completed Faecal Immunochemical Test (FIT). A FIT is a test that checks for occult (hidden) blood in the stool. The test detects globin levels in faeces and yields a score. **BowelScreen's score is measured in *haemoglobin nanograms per millilitre* (ngHb/ml)** To determine whether a FIT is positive, the FIT score is measured against a pre-set "cut-off score". The higher the cut-off score, the lower the positivity rate and the higher the specificity and positive predictive value (PPV) for colorectal cancer (CRC) and advanced adenomas. **Current cut-off:** 225 haemoglobin nanograms per millilitre (225ngHb/ml). **Positive FIT:** The FIT is deemed positive if the measured amount of blood (FIT result score) is above the cut-off level (cut-off score). **Negative FIT:** The FIT is deemed negative if the measured amount of blood (FIT result score) is equal to or below the cut-off level (cut-off score).

Different types of unit measurement	Conversion Guide
BowelScreen FIT score data is captured in nanograms of haemoglobin per millilitre (ngHb/ml). However, other bowel screening programmes can capture FIT scores in micrograms of haemoglobin per gram of faeces (µgHb/g).	1 µgHb/g = 5 ngHb/ml 1 ngHb/ml = 0.2 µgHb/g **225 ngHb/ml = 45 µgHb/g** **100 ngHb/ml = 20 µgHb/g**

Historical changes to the FIT "cut-off score"		
Period	**BowelScreen cut-off score**	**Clinical Advisory Group (CAG) date**
Start to 03/02/2014	100 nanograms of haemoglobin per millilitre (100ngHb/ml) **or** 20 micrograms of haemoglobin per gram (20µgHb/g)	16/07/2012
03/02/2014 to present	225 nanograms of haemoglobin per millilitre (225ngHb/ml) **or** 45 micrograms of haemoglobin per gram (45µgHb/g)	03/02/2014

Source: cut-off changes also reported in *A National Bowel Cancer Screening Programme using FIT: Achievements & Challenges*[25]

Historical changes to the BowelScreen eligible age range	
Period	**Eligible Age Range**
Programme Start - October 2023	60 to 69 years old
October 2023 - April 2025	59 to 69 years old
April 2025 - October 2025	59 to 70 years old
October 2025 - Present	58 to 70 years old

Appendix D – Abbreviations and measurements

Abbreviations	
ACS	American Cancer Society
AJCC	American Joint Commission on Cancer
BSP	Bowel Screening Programme
CAG	Clinical Advisory Group
CDR	Cancer Detection Rate
C-RADS	CT Colonography Reporting And Data System
CRC	Colorectal Cancer
CSO	Central Statistics Office
CTC	Computerised Tomography Colonography
DSP	Department of Social Protection
DICOM	Digital Imaging and Communications in Medicine
DOH	Department of Health
DQM	Data Quality Manager
DQS	Data Quality Statement
ERG	Expert Reference Group
ERS	Endoscopy Reporting System
EU	European Union
FIT	Faecal Immunochemical Test
GDPR	General Data Protection Regulations
GI	Gastrointestinal
GP	General Practitioner
Hb	Haemoglobin
HIQA	Health Information & Quality Authority
HL7	Health Level 7
HL7 FHIR	HL7 - Fast Healthcare Interoperability Resources
HSE	Health Service Executive
HTA	Health Technology Assessment
IACP	Irish Association of Coloproctology
ICCR	International Collaboration on Cancer Reporting
ICD-10	International Classification of Disease 10th Edition
ICR	Interval Cancer Rate
IHI	Individual Health Identifier
ISO	International Standards Organisation
IT	Information Technology
KPI	Key Performance Indicator
LOINC	Logical Observation Identifiers, Names & Codes
MDM	Multidisciplinary Meeting
NCCP	National Cancer Control Programme
NCRI	National Cancer Registry Ireland
NSS	National Screening Service
NVARCHAR	National Varying Character
PCCRC	Post-Colonoscopy Colorectal Cancer
PDF	Portable Document Format
PEU	Programme Evaluation Unit
PPV	Positive Predictive Value
QA	Quality Assurance
RCPath	Royal College of Pathologists
RCPI	Royal College of Physicians Ireland
RCSI	Royal College of Surgeons Ireland
SSRS	SQL Server Reporting Services
SNOMED	Systemised Nomenclature of Medicine
SQL	Structured Query Language
TNM	Tumour Node Metastasis
UICC	Union for International Cancer Control
VARCHAR	Variable Character
WHO	World Health Organization

Measurements	
ng	nanogram
Hb/ml	haemoglobin per millilitre
ngHb/ml	nanograms of haemoglobin per millilitre
µg	microgram
Hb/g	haemoglobin per gram
µgHb/g	micrograms of haemoglobin per gram

Appendix E – Definitions

Term	Definition
Boolean data	*Boolean* is a data type which has one of two possible values (e.g., true or false). Boolean data is dichotomous data.
Cut-off score	The *'cut-off' score* is the score above which a faecal immunochemical test (FIT) is deemed positive (abnormal). The 'cut-off' score is also known as the *'threshold score'*.
Data controller	*Data controller* is any person, authority or agency that determines the purpose and manner of personal data processing[28].
Data processor	A *Data processor* refers to a person, company, or other body which processes personal data on behalf of a data controller. HSE guidance states that this does not include an employee of the controller who processes data during the course of their employment[28].
Data	*Data* are the building blocks for information. Data may be described as numbers, symbols, words, images and graphics that have been validated but yet to be organised or analysed[14].
Database	A *database* is a collection of data that is organised so that its contents can easily be accessed, managed, and updated[14].
Data dictionary	A *data dictionary* is a descriptive list of names, definitions, and attributes of variables collected in a database or dataset[14].
Data linking	*Data linking* (*syn. Data Linkage*) is the process of collating data from different sources to create a more valuable & helpful dataset.
Data protection impact assessment	A *Data protection impact assessment* is a process designed to identify risks arising out of the processing of personal data and to minimise these risks as far and as early as possible[14].
Data quality	*Quality data* is data which is *'fit for purpose'*. The quality of data is determined through assessment against five internationally accepted data quality dimensions[14].
Data quality dimensions	*Data quality dimensions* are a set of data quality attributes that represent a single aspect or construct of data quality. These dimensions are relevance; accuracy & reliability; timeliness & punctuality; coherence & comparability; and accessibility & clarity[14].
Data subject	A *data subject* is any identified or identifiable living individual to whom personal data relates[6].
Data users	*Data users* are any stakeholders that use data held by a screening programme or use data outputs from that programme[3].
Event-driven application	An *event-driven application* is a database designed to respond to actions generated by the user (or the system) to produce, detect or react to events. An event is defined as a *'change in state'*.
Faecal immunochemical test (FIT)	The *faecal immunochemical test* (FIT) is a screening test for colon cancer. A small quantity of stool is collected using a sample stick and sent to the laboratory for analysis. The test looks for hidden blood in the stool, which can be an early sign of cancer.
Health information standards	*Health information standards* is a broad term which encompasses three distinct categories of data standards for health data, namely: *messaging standards (e.g., HL7, DICOM)*, *data dictionaries*, and *terminology & classification standards (e.g., SNOMED-CT, LOINC, ICD-10)*[24].
Individual health identifier (IHI)	An *Individual health identifier* is a number that uniquely & safely identifies each person that has used, is using or may use Irish health services[22].
Information	*Information* is data that has been processed or analysed to produce something useful[14].
Information management standards	*Information management standards* are a set of standards outlined by Ireland's Health Information and Quality Authority (HIQA), designed to provide national health and social care organisations with structure and guidance in the management of the data and information collected, held, and disseminated by those organisations. In 2018, HIQA outlined 10 standards across 6 themes. In 2024, HIQA outlined 12 standards across 3 principles.
International classification of disease (ICD)	*International classification of disease* (ICD) is a type of *terminology & classification standard*. ICD is a globally utilised disease classification system. It provides a system of codes for classifying diseases, signs, symptoms, abnormal findings, complaints, social circumstances, and external causes of injury or disease[24]. It is owned, developed, published and updated by the WHO.
Interval cancer	A *post-colonoscopy colorectal cancer* (PCCRC) is the diagnosis of a colorectal cancer (CRC) within 36 months of last BowelScreen colonoscopy, that was reported as a negative colonoscopy where no abnormality was found, or a polyp was removed.

Term	Definition
Key performance indicators	*Key performance indicators* are specific and measurable elements of practice that are designed to assess key aspects of structures, processes, and outcomes[3].
Messaging standards	*Messaging standards* specify the structure & order of the elements that form a message exchanged between IT systems (e.g., HL7, HL7 FHIR, DICOM)[24].
Metadata	*Metadata* is 'data to explain data'. It provides structured information about the content of a dataset or database (e.g., data dictionaries, data quality statements, data profiles, etc.)[26].
Modified Gloucester comfort score	*Modified Gloucester comfort score* is a five-point scale for measuring patient comfort during colonoscopy. It considers the frequency and duration of discomfort and any distress it might cause the patient.
Paris classification	The *Paris classification* is an international standard for endoscopic classification of gastrointestinal superficial neoplastic lesions, adapted from the Japanese macroscopic classification for gastric cancer.
Portable document format (PDF)	*PDF* is a digitally accessible, non-proprietary, non-machine-readable format for documents and information. Under the open data initiative, PDF is a one-star open data format.
Positive predictive value	The *positive predictive value* of a screening test is the probability that an individual who had a positive screening test actually has the disease[27].
Privacy statement	A *Privacy statement (*sometimes called a *statement of information practices)* is a publicly available document that sets out what information is collected; how it's used; with whom it is shared & for what purpose; the protection safeguards in place; and how service users can access their information[14].
Record matching	*Record matching* (*syn. Data Matching*) is the process of identifying whether two records in a database refer to the same real-world entity.
Screening register	A *population-based screening register* (or population register) holds data on the target population in geographically defined area[27].
Semantic interoperability	*Semantic interoperability* is the ability of health IT systems to exchange/interpret data, and actively use exchanged data. Semantic Interoperability requires shared syntax & semantics[24].
SNOMED-CT	*Systematized nomenclature of medicine-clinical terms* (SNOMED-CT) is a type of *terminology & classification standard*. SNOMED-CT is managed by Ireland's National Release Centre (NRC). It is a systematically organized computer-processable collection of multilingual medical terms providing codes, terms, synonyms, and definitions used in clinical documentation and reporting[24].
Special category data	*Special category data* is personal data revealing racial or ethnic origin, political opinions, sex life or orientation, religious or philosophical beliefs, or trade union membership, genetic data, biometric data which uniquely identifies a natural person, or data concerning health[9].
Specificity	*Specificity* is the ability of the screening test to identify healthy people as negative for a particular disease. It is proportion of individuals without a certain disease who were identified as non-diseased by the screening test[27].
SQL server reporting services	*SQL server reporting services* (SSRS) is a SQL Server subsystem that enables the creation of graphical, mobile, and printed reports using SQL Server and other data sources.
Statement of purpose	A *statement of purpose* is a publicly available document which outlines why the data collection exists & the objectives of the data collection[14].
Terminology & classification standards	*Terminology & classification standards* assign a unique code or value to clinical concepts such as diseases, medications & procedures. Examples include ICD-10, SNOMED-CT, and LOINC[14].
TNM staging	The *TNM classification of malignant tumours* (TNM) is an internationally recognised standard for classifying the extent of spread of cancer. It is a classification system of the anatomical extent of tumour cancers. It is sometimes referred to as the *AJCC pathological staging system*.